Mutation-selection models of sequence evolution in population genetics

Thesis

Unrestricted.

Mutation–selection models of sequence evolution in population genetics

by

Dipl.-Phys. Tini Garske

A dissertation submitted for the degree
Doctor of Philosophy
in
Applied Mathematics,
Faculty of Mathematics and Computing,
The Open University

November 2004

Submission date: 16 November 2004
Award date: 22 March 2005.

ABSTRACT OF THE DISSERTATION

Mutation–selection models of sequence evolution in population genetics

by

Dipl.-Phys. Tini Garske

Doctor of Philosophy in Applied Mathematics,
Faculty of Mathematics and Computing, The Open University, 2004

The equilibrium properties of a number of deterministic mutation–selection models of sequence evolution are investigated. Both two- and four-state sequences are considered, the mutation model is a single-step mutation model. Two types of fitness functions are studied, namely permutation-invariant fitness functions, where the fitness of a sequence depends only on the number of mutations, not on their location within the sequence, and Hopfield-type fitness functions, where the fitness of a sequence is determined by its similarity to a number of predefined patterns.

Maximum principles to determine the population mean fitness in equilibrium are derived, where the maximiser gives also the ancestral mean genotype. These maximum principles are used to investigate the error threshold phenomenon, i.e., the phenomenon that for certain fitness functions the population changes at a critical mutation rate from a well localised to a delocalised distribution in sequence space.

The error threshold phenomenon is investigated for a four-state model with permutation-invariant fitness functions and for a two-state model with Hopfield-type fitness functions. Both models yield ordered and disordered as well as partially ordered phases.

Acknowledgments

I would like to thank my supervisor Uwe Grimm for help, advice and all the things I've learned from him during the past three years. Furthermore, I wish to thank Michael Baake, Ellen Baake and Oliver Redner for advice and insightful discussions, helpful feedback on my work and valuable comments on various manuscripts.

I also wish to gratefully acknowledge the Erwin Schrödinger Institute in Vienna for support during my visits in December 2002 and December 2003, providing me with the opportunity to work with a wonderful group, and the British Council for funding my travel to collaborate with the research group of Michael Baake in Bielefeld.

Last but not least a big thank you to my partner Kit and my parents for their support, encouragement and unfailing patience.

Publications

The derivation of a maximum principle for a four-state model with permutation-invariant fitness function that is obtained in section 4.1.2 and chapter 5 (specialised to the permutation-invariant four-state case) has been published in [GG04b]. The investigation of the error threshold phenomenon for a four-state model with permutation-invariant fitness functions as shown in section 6.1 has been published in [GG04a].

References

[GG04b] T. Garske and U. Grimm. A maximum principle for the mutation–selection equilibrium of nucleotide sequences. *Bulletin of Mathematical Biology*, 66(3):397–421, 2004.

[GG04a] T. Garske and U. Grimm. Maximum principle and mutation thresholds for four-letter sequence evolution. *Journal of Statistical Mechanics: Theory and Experiment*, P07007, 2004.

TABLE OF CONTENTS

List of Figures

List of Tables

CHAPTER 1

Introduction

Population genetics is concerned with the investigation of the genetic structure of populations. This structure is influenced by evolutionary factors such as mutation, selection, recombination, migration, and genetic drift. Hence the investigation of interactions between these processes lies at the heart of research in population genetics. For an excellent review of the theoretical aspects of this field see [BG00] or [CK70, Ewe04].

Treating biological evolution theoretically involves, as a first step, the configuration of a model, i.e., abstracting from reality and capturing only the essential features in a mathematically tractable formulation. Incorporating all of the evolutionary factors mentioned above into the model leads to a model that is too complex to be of practical use. One approach to further the understanding of evolution is therefore to use simpler models that restrict the number of evolutionary factors to a subset, excluding others. In this thesis the antagonistic interplay of mutation and selection shall be investigated, with mutation generating the genetic variation upon which selection can act, and thereby tending to reduce it.

Pure mutation–selection models exclude genetic drift and therefore are deterministic models, and accurate only in the limit of an infinite population size; for a review see [Bür00]. A further simplification taken here is to consider only *haploid* populations, i.e., asexually reproducing populations, where the genetic code exists in one copy only in each cell, in contrast to *diploid* populations that reproduce sexually, with each individual having two copies of the genome, one stemming from the mother and one from the father. However, if one considers a diploid population neglecting the effects of dominance, the resulting equations are actually the same as in the haploid case [Bür00].

The evolution of a population under the forces of mutation and selection can be described mathematically by a set of differential equations, the evolution equations. If one is interested in equilibrium properties, the problem reduces to an eigenvalue problem of a matrix of dimension given by the number of genotypes allowed for in the model.

In the early works on population genetics (e.g. [Fis30, Hal27]), the concept of *genes* or *genotypes* was a rather abstract one. With the discovery of the structure of the DNA [WC53], however, genotypes could be identified with sequences written in the four-letter alphabet AGCT (denominating the four nitrogenous bases adenine, guanine, cytosine and thymine), which is the basis of the *sequence space approach*, introduced in the seminal paper by Eigen [Eig71].

This class of models gets its name from the sequence space which is the set of all possible sequences and thus genotypes. As there are 4^N different sequences of a given length N, the number of possible genotypes is very large, especially compared to classical mutation–selection models, where in the simplest cases only two different types are considered. This large number of types makes the mathematical treatment of the evolution equation enormously harder as the dimension of the problem grows with the number of types considered.

Most works using the sequence space approach do not use the full four-letter alphabet, but a simplified two-letter version, which can for instance be interpreted as purines and pyrimidines. Results concerning the full four-state model are scarce, but those that are available show a considerably richer behaviour.

In the sequence space approach, the model is based on the microscopic level, at which mutation occurs. Consequently, mutation is straightforward to model. However, the appropriate modelling of selection is a more challenging task, as selection acts on the phenotype, and the mapping from sequence to phenotype is by no means simple. To this end, the concept of the fitness landscape [KL87] is introduced as a function on the sequence space, assigning a fitness value, which determines the reproduction rate, to each possible genotype. Apart from the problem that a realistic fitness landscape would have to

be highly complex (too complex for a mathematical treatment) there is also very limited information available concerning the nature of realistic fitness functions.

Since the 1980's, one aspect driving the scientific progress has been the discovery of the equivalence of some types of sequence space models with certain models used in statistical physics, such as Ising models and quantum spin chains [Leu86, BBW97]. This has made the methods from theoretical physics, that are well adapted to dealing with large numbers of particles (or types), available for use in population genetics. However, the application of the methods from physics to biology is not completely straightforward due to subtle differences in the normalisation of the relevant observables.

One result that was obtained influenced by methods from statistical physics is a simple scalar maximum principle to determine the population mean fitness in equilibrium for a two-state model with *permutation-invariant* fitness function, where the fitness of a sequence depends only on the number of mutations it carries compared to a reference sequence, not on their location within the sequence. This is an immense simplification compared to the solution of a high-dimensional eigenvalue problem which previously was required to find the population mean fitness in equilibrium. The analogon of this maximum principle is the principle of minimal free energy in statistical physics.

The general behaviour of a deterministic mutation–selection model is such that, for a fixed fitness function, at low mutation rates the population will be clustered around the highest peak of the fitness landscape, and with increasing mutation rate this cluster will broaden, until in the limit of infinite mutation rates, the population will be equally distributed over the whole sequence space, as selection can no longer counteract the dispersing effect of mutation. Using the sequence space approach, this behaviour shows however an interesting twist: While for some (predominantly simple, smooth) fitness landscapes, the broadening of the population distribution occurs in a smooth fashion, there are other fitness functions for which the population distribution shows a radical change at a critical mutation rate. This phenomenon is known as the *error threshold*, and has attracted considerable interest ever since it has first been observed [Eig71]. It has been long de-

bated whether this was just an artefact of the model employed or whether it described a behaviour that is relevant to real populations. Today, there exists however ample experimental evidence for the occurrence of error thresholds in viral populations [DH97, CCA01].

This thesis is organised as follows. Chapters 2 and 3 give an introduction to the models employed in this work. Chapter 2 gives the general set-up of the mutation–selection model. Chapter 3 focuses on the sequence space approach and the adaptations of the general model to this approach, specifying the particular mutation models for both two- and four-state sequences and discussing some types of fitness functions, in particular the permutation-invariant fitness and the *Hopfield-type fitness*, where the fitness of a sequence is determined by the similarity to a number of given sequences.

In chapter 4 the problem of the large number of genotypes that is inherent in the sequence space approach is addressed by lumping together types with identical fitness values into classes, and considering a coarser process that keeps track of the classes, rather than individual types. Furthermore, this lumping is an essential step toward the derivation of the maximum principle, which is pursued in chapter 5. Both lumping and maximum principle are obtained for two- and four-state models with permutation-invariant and Hopfield-type fitness functions.

Chapter 6 uses the maximum principles derived in chapter 5 to investigate the phenomenon of the error threshold for examples of permutation-invariant and Hopfield-type fitness functions.

The results of this thesis are summarised in chapter 7.

CHAPTER 2

General mutation–selection models

In a deterministic mutation–selection model, the only evolutionary forces that are considered are mutation and selection. Usually, one is interested in questions concerning the mutation–selection balance, means and variances and the mutation load, i.e., the loss of population mean fitness due to the action of mutation.

The subject under investigation is a population, which can be considered as a set of individuals. These individuals have a type, and in the case of a pure mutation–selection model, this type is identified with the genotype, neglecting any environmental effects. The set of possible types is referred to as the *type space* $\mathfrak{S}$. This type space can be finite; in the simplest case one considers only one gene or *locus* that occurs in two different forms, called *alleles*, and the type space $\mathfrak{S}$ contains only two types, having cardinality $|\mathfrak{S}| = 2$. However, the type space $\mathfrak{S}$ can as well be infinite as in the continuum-of-alleles model, describing the possible genotypes by real numbers, see for instance [Bür00, Red03].

Whatever the type space is, as long as it is finite, the population at any time t can be described as a vector $\boldsymbol{p}(t)$ of dimension $|\mathfrak{S}|$. An entry $p_i(t)$ determines the fraction of the population that is of type i. The population is normalised such that $\sum_{i\in\mathfrak{S}} p_i(t) = 1$ and therefore is a probability distribution. Alternatively, the population can be described by the numbers $n_i(t)$ of each type, collected in the vector $\boldsymbol{n}(t)$, which is not normalised.

In the setting of a deterministic, haploid mutation–selection model (or a diploid model without dominance effects), there are various ways to model the evolutionary processes. One way is to consider time as a discrete quantity, that is measured in (non-overlapping) generations, where mutations occur at reproduction. An example for this approach can be found in [Leu86].

If one decides to model time as a continuous quantity, mutation can still be modelled as coupled to the reproduction process as it is done in the original quasispecies model [Eig71], or, alternatively, as a process that occurs independently from reproduction. Models following the latter approach have been termed *para-muse models* (*para*llel *mu*tation-*se*lection models) [Baa95]. The mutation–selection equation for the parallel model has been given in [CK70] and investigated in [Aki79, Wag98, BW01] and others.

For organisms with short replication periods, like bacteria or viruses, it makes sense to assume that most of the mutations occur as replication errors and therefore the coupled model seems to be more realistic. For organisms with longer generations, however, replication errors are not the only source of mutations, but mutations also occur during the lifetime of an individual, a process that is correctly modelled by the decoupled version. Which of these one deems to be more realistic depends on whether one considers mutation rates to be constant per generation or per year. The latter is assumed for the molecular clock hypothesis, which forms a basis of some important approaches in fitting molecular phylogenies. The very existence of the molecular clock has been challenged [Gil91], but studies investigating the generation-time effect on the molecular clock only measure moderate dependence on generation length [Kim87, Oth93]. As so often, reality lies somewhere in between these two extremes.

In [WBS95], both versions, the coupled and decoupled mutation–selection schemes, have been compared. For small mutation rates, they are equivalent in first approximation, and even for large mutation rates, both models were found to behave very similarly. The mathematically simpler model is the para-muse model, which will be the focus in this work.

In this model, the evolutionary processes are births, deaths and mutations. The birth process describes the event that one individual in the population gives rise to an identical copy of itself. In the case of a bacterial population, this would be a cell division. The death process describes the event of an individual dying and therefore vanishing from the population. The births and deaths happen with rate b_i and d_i, respectively, which depend

on the type i. Taken together, these two processes determine the effective reproduction rate $r_i = b_i - d_i$. The type-dependence of the reproductive process is how selection enters this model: A type with high effective reproduction rate, or fitness, has a selective advantage.

A mutation event is a type change of an individual, which happens with a rate $m_{ji} (= m_{j \leftarrow i})$, depending on both the initial type i and final type j. At this level, the modelling of the population is straightforward, the meat lies in the choice of values for the mutation and reproduction rates.

Both reproduction and mutation rates can conveniently be collected in square matrices $\mathcal{R}$ and $\mathcal{M}$. The reproduction matrix $\mathcal{R}$ is diagonal with entries r_i, whereas the off-diagonal entries of the mutation matrix $\mathcal{M}$ contain the mutation rates. As the pure mutation process does not change the number of individuals, the diagonal entries are chosen such that $\mathcal{M}_{ii} = -\sum_{\substack{j \in \mathfrak{S} \\ j \neq i}} \mathcal{M}_{ji}$, specifying the rate at which individuals mutate away from type i to any other type. This makes the mutation matrix $\mathcal{M}$ a Markov generator. The time-evolution operator $\mathcal{H}$ is given by the sum of reproduction and mutation matrix, $\mathcal{H} = \mathcal{R} + \mathcal{M}$.

In the deterministic limit, the evolution of the population is governed by a differential equation

$$\dot{\boldsymbol{p}}(t) = [\mathcal{H} - \bar{r}(t)\mathbf{1}]\, \boldsymbol{p}(t) \,. \tag{2.1}$$

Here, $\bar{r}(t)$ is the population mean fitness, defined as $\bar{r}(t) = \sum_i r_i p_i(t)$, and $\mathbf{1}$ is the identity matrix. For the connection to the corresponding stochastic model, see section 2.1, and equation (2.8) in particular. Note that equation (2.1) is non-linear, due to the population mean fitness, which depends on $\boldsymbol{p}(t)$. This term is required to preserve the normalisation of the population distribution $\sum_i p_i(t) = 1$.

The non-linearity of equation (2.1) can easily be remedied by using a transformation from the frequencies $\boldsymbol{p}(t)$ to absolute numbers of individuals, normalised by the population

size at time $t = 0$, $\boldsymbol{y}(t) = \frac{\boldsymbol{n}(t)}{\sum_i n_i(0)}$. One gets [TM74]

$$\boldsymbol{y}(t) := \boldsymbol{p}(t) \exp\left[\int_0^t \bar{r}(\tau) d\tau\right] , \tag{2.2}$$

with reverse transformation

$$\boldsymbol{p}(t) = \frac{\boldsymbol{y}(t)}{\sum_i y_i(t)} . \tag{2.3}$$

$\boldsymbol{y}(t)$ now fulfils the linearised evolution equation

$$\dot{\boldsymbol{y}}(t) = \mathcal{H}\boldsymbol{y}(t) . \tag{2.4}$$

Defining

$$\boldsymbol{T}(t) := \exp[t\mathcal{H}] , \tag{2.5}$$

which is given via the Taylor expansion $\exp[t\mathcal{H}] = \sum_{n=0}^{\infty} \frac{t^n}{n!}\mathcal{H}^n$, the solution of this equation is clearly given by

$$\boldsymbol{y}(t) = \boldsymbol{T}(t)\boldsymbol{y}(0) , \tag{2.6}$$

and thus we have for the frequencies

$$\boldsymbol{p}(t) = \frac{\boldsymbol{T}(t)\boldsymbol{p}(0)}{\sum_{ij} T_{ij}(t) p_j(0)} . \tag{2.7}$$

The main interest focuses on the equilibrium, i.e., the behaviour if $\dot{\boldsymbol{p}} = 0$, which is attained for $t \to \infty$. All equilibrium quantities shall be denoted by omitting the argument t, for instance $\boldsymbol{p}$ is the equilibrium population distribution and $\bar{r}$ the equilibrium mean fitness. In equilibrium, equation (2.1) reduces to an eigenvalue equation for $\mathcal{H}$ with leading eigenvalue $\lambda_{\max} = \bar{r}$. If $\mathcal{M}$ is irreducible, as shall be assumed throughout (unless otherwise specified), Perron-Frobenius theory ([Per07], [Kar66, appendix]) applies, which guarantees that the leading eigenvalue $\bar{r}$ of $\mathcal{H}$ is non-degenerate, and the corresponding right eigenvector $\boldsymbol{p}$ is strictly positive, which implies that it can be normalised as a probability distribution.

The continuous-time model is connected to a discrete-time mutation–selection model with a time-evolution operator given by $\boldsymbol{T}(t)$, where t is the generation time. The pure mutation process in continuous time is a Markov process with $\sum_j \mathcal{M}_{ji} = 0$, whereas the

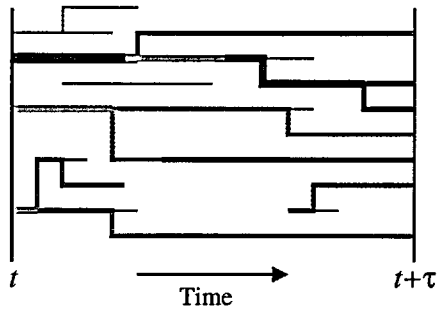

Figure 2.1: Schematic population: lineages symbolise individuals, colours indicate types. Colour changes mark mutation events, whereas branching lines represent births and ending lines stand for deaths. The thickness of a line at any time denotes the number of offspring of the individual at time $t + \tau$.

pure mutation process in discrete time is given by the corresponding Markov chain with $\sum_j (\exp[t\mathcal{M}])_{ji} = 1$.

2.1 Relative reproductive success and the ancestral distribution

Similarly to the population distribution, there is another distribution that will prove important in this model, namely the *ancestral distribution*, which has been introduced in [HRWB02] via considering a stochastic model that corresponds to the deterministic model in question. The stochastic model is a multi-type branching process in continuous time with birth, death and mutation events with rates b_i, d_i and m_{ji}, respectively. As in the deterministic model, mutation and reproduction are modelled as decoupled, the multi-type branching process in the corresponding stochastic model is such that mutations never occur at birth events, i.e., the offspring is always of the same type as the parent. Figure 2.1 shows schematically the evolution of a population according to such a branching process.

This stochastic model corresponds to the deterministic model in the following sense. In the limit of an infinite population size, the numbers of individuals $\boldsymbol{Y}(t)$, which are random variables in the stochastic model, divided by the initial population size, converge almost surely to the solution $\boldsymbol{y}(t)$ of equation (2.4), see [EK86, chapter 11]. Also, for finite population size, the time evolution of the expectation values $E[\boldsymbol{Y}(t)|\boldsymbol{y}(0)]$ (with some

initial population $\boldsymbol{y}(0)$) is governed by

$$\frac{d}{dt}E[\boldsymbol{Y}(t)|\boldsymbol{y}(0)] = \mathcal{H}E[\boldsymbol{Y}(t)|\boldsymbol{y}(0)]\,, \tag{2.8}$$

such that $\mathcal{H}$ is the infinitesimal generator of this process [HRWB02].

In particular, consider a population that was started at time 0 by a single type-j individual. Clearly, the expected number of type-i offspring at time τ is given by $T_{ij}(\tau)$ with $\boldsymbol{T}$ from (2.5). Due to independence of individuals, this is also the expected number of type-i offspring at time $t+\tau$ of a type-j individual at time t that is part of a population distributed according to $\boldsymbol{p}(t)$. Note that due to the Markov property of the process, this depends only on the time increment τ, not on the time of founding t. The expected total number of offspring of this individual, i.e., the expected *clone size*, after evolution for time period τ is then $\sum_i T_{ij}(\tau)$.

Comparing the expected clone size of a j-clone after evolution for time τ with the expected mean clone size of the population, distributed according to $\boldsymbol{p}(t)$, $\sum_{k\ell} T_{k\ell}(\tau)p_\ell(t)$, the relative reproductive success $z_j(\tau,t)$ is given by

$$z_j(\tau,t) := \frac{\sum_i T_{ij}(\tau)}{\sum_{k\ell} T_{k\ell}(\tau)p_\ell(t)}\,. \tag{2.9}$$

The concept of the ancestral distribution was introduced in the context of sequence space models in [HRWB02], and, maybe more importantly, recognised for its relevance, as shall be seen later. However, ideas using the concept of the relative reproductive success can already be found in [Fis30], and they have been used in matrix population models. These models are formally very similar to the sequence space models, but the types are not interpreted as genotypes but rather as age stages in a plant population, and the change from one class to another is not due to mutation but due to growth, see [CE92] or [Cas01, chapter 4.6].

Picking at random an individual at time $t+\tau$ and tracing its ancestry back in time until t, we are interested in the probability distribution of the ancestral type at time t, irrespective of the type at time $t+\tau$. This is denoted by the ancestral distribution $\boldsymbol{a}(\tau,t)$. As the relative reproductive success of type j, $z_j(\tau,t)$, determines the frequency of the

lines at $t+\tau$, which originate from *each* type-j individual at t, the ancestral distribution $\boldsymbol{a}(\tau,t)$ is given by

$$a_j(\tau,t) := z_j(\tau,t)p_j(t)\ . \tag{2.10}$$

The ancestral distribution can not only be interpreted backward in time, as explained above, but also admits an interpretation in forward time as the population distribution at time t, weighted by the number of offspring each type produces at time $t+\tau$ (irrespective of the offspring's type).

2.1.1 Asymptotics

It is well known (e.g. [Kar66, appendix]), that in the limit $t \to \infty$ the operator $\exp[t(\mathcal{H}-\lambda_{\max}\mathbf{1})]$ converges to a matrix $\boldsymbol{P}$ with $P_{ij} = p_i z_j$, where p_i and z_j are elements of the right and left Perron-Frobenius (PF) eigenvectors, respectively, normalised such that $\boldsymbol{zp} = 1$, and that $\boldsymbol{P}$ is a projector onto $\boldsymbol{p}$.

From this and the definition of the relative reproductive success (2.9), it can be seen that in equilibrium ($t, \tau \to \infty$) the $z_j(\tau,t)$ converge to the elements of the left PF eigenvector $\boldsymbol{z}$ of $\mathcal{H}$, using the identity $\exp[t(\mathcal{H}-\lambda_{\max}\mathbf{1})] = e^{-\lambda_{\max}}\boldsymbol{T}(t)$. The ancestral frequency of type i in equilibrium is then given by $a_i = z_i p_i$, and is normalised as a probability distribution, i.e., $\sum_i a_i = 1$.

2.2 Means

In this model, there are two important probability distributions, the population and ancestral distributions. Hence, any property (for instance, the fitness) can be averaged with respect to either distribution. The *population mean* of a quantity o, that is assigned to each type in the sequence space, is given by

$$\bar{o}(t) = \sum_i o_i p_i(t)\ , \tag{2.11}$$

whereas the *ancestral mean* is

$$\hat{o}(\tau, t) = \sum_i o_i a_i(\tau, t) \ . \tag{2.12}$$

Note that the value of o_i for any type i is taken to be constant in time; the time dependence of the means only comes in through the time dependent distributions. The values of the means in equilibrium shall be denoted as before by omitting the time arguments t and τ.

The difference between the population mean and the ancestral mean is rather subtle: Both the population and ancestral distributions are normalised as probability distributions, $\sum_i p_i = \sum_i a_i = 1$. However, in terms of the eigenvectors of the time-evolution matrix the population mean can be thought of as a mean with respect to an L_1-norm, because in the mean of any function the functional values are weighted linearly with the entries of the eigenvectors. This is different for the ancestral distribution, as it is given by the product of the left and right eigenvectors of the time-evolution matrix $a_i = z_i p_i$, which, after symmetrisation of the time-evolution matrix, corresponds to the square of the eigenvectors, as then the left and right eigenvectors coincide (cf. chapter 5). This means that in the ancestral mean the functional values are weighted with the square of the entries of the eigenvectors, and thus this can be interpreted as a mean with respect to an L_2-norm.

This difference has made the application of methods from physics to biology more difficult, as in the biological models, the natural observables are the population means, whereas in the physical models, one works with the L_2-norms such as the ancestral mean. However, it has turned out that also in biology, the ancestral mean plays an important role, as shall be seen in chapters 5 and 6.

CHAPTER 3

Mutation–selection models in the sequence space approach

3.1 Sequences as genotypes

In the previous chapter, a general haploid mutation–selection model in continuous time was set up without any restrictions of the types considered, apart from the fact that they were assumed to be genotypes of some form, i.e., inheritable and not influenced by the environment. Here, the notion of types shall be filled with a particular concept. Genotypes are modelled as sequences inspired by the structure of the DNA and RNA.

The DNA and RNA are sequences composed of four different nucleotides, which differ in the nitrogenous base. In DNA, we find adenine, guanine, cytosine and thymine (AGCT), whereas in the RNA, uracil (U) replaces thymine. These nitrogenous bases can be classified as purines (AG) and pyrimidines (CTU). The genetic information stored in the DNA is given by the succession of the different nucleotides.

The simplest approach is to model the sequential structure of the DNA, but neglect the full complexity of the four-letter alphabet, by modelling genotypes as two-state sequences. This is the oldest sequence space approach [Eig71], and up to date by far the most common in population genetics, see [NS89, FPS93, BBN96, LTP03], to name but a few. Here, these two states can be identified as purines and pyrimidines. Alternatively, one can identify one reference sequence, commonly chosen as the *wild-type*, which is the sequence with highest fitness and the one usually found most abundant in natural populations. Then, the two states can be taken to indicate whether a given sequence matches the wild-type at a particular site or not.

The more complex approach to allow for four different states at each site as it occurs in

actual DNA has only been considered in very few works to date [HWB01, GG04b, GG04a]. The results of these studies show however a considerably more complex structure that is worthwhile investigating.

Although the two- and four-letter alphabets are the natural choice to model the structure of DNA, the sequence space approach is by no means limited to these alphabets. One could just as well consider proteins, which are sequences of amino acids, see [WG01, WLG01]. Although the alphabet size is increased to 20 in this case, the number of types actually decreases, because an amino acid sequence has only one third of the length of the corresponding nucleotide sequence.

Mathematically spoken, consider sequences of fixed length N written in an alphabet $\mathcal{A}$. For two-state models the alphabet is commonly chosen as $\mathcal{A} = \{+, -\}$, but $\mathcal{A} = \{0, 1\}$ is useful in some instances as well. In four-state models, one uses $\mathcal{A} = \{\mathtt{A}, \mathtt{G}, \mathtt{C}, \mathtt{T}\}$, $\mathcal{A} = \{++, +-, -+, --\}$, or $\mathcal{A} = \{0, 1, 2, 3\}$. Each sequence $\boldsymbol{\sigma} \in \mathfrak{S}$ has N sites $\sigma_\alpha \in \mathcal{A}$, $\alpha \in \{1, \ldots, N\}$. Hence the sequence space is the power set of the alphabet, $\mathfrak{S} = \mathcal{A}^N$, with the number of different sequences $|\mathfrak{S}| = |\mathcal{A}|^N$.

If one generalises the alphabet such that sites are not identified with nucleotides, but with general *loci*, i.e., any genes that contribute to the character in question, the sequence space models are equivalent to multilocus models with *complete linkage*, i.e., the case of a multilocus model evolving free from recombination. The multilocus and sequence space models have been developed independently from each other, and their equivalence has only been noticed later [Hig94], for a discussion see [BW01].

3.2 Mutation models

The choice of mutation model defines a "neighbourhood" within the sequence space [RS02], and therefore a structure is imposed on the sequence space. Sequences that can mutate into one another in one mutational step are considered to be neighbours.

In a rather simple form, the modelling of mutation is straightforward in the sequence

space approach. Disregarding more complex mechanisms that change the length of a sequence, like deletions, insertions or duplications, only single point mutations are taken into account, i.e., replacements of one nucleotide by another. This is known as the *single step mutation model*, introduced in [OK73]. Of course this is a severe simplification, as insertions or deletions in coding regions can shift the reading frame and therefore a single mutation of that type may render the whole sequence infunctional. Furthermore, duplications are known to occur in evolution, and are thought to be one mechanism how new functions evolve: After a duplication of a gene, there is a spare copy of that gene, which evolves free from purifying selection and can therefore "try out" different other versions without loosing the initial functionality which is maintained by the original copy of the gene. However, in the framework of this model, even the simplified single step mutation model is better justified than the fitness functions that are currently tractable.

3.2.1 Two-state models

Assuming equal mutation rates at each site, which clearly is a simplification of reality [Yan96, CL04], for a two state model there are two possible mutation rates, μ^+ governing the process $+ \to -$ and μ^- governing $- \to +$. If the two states are taken to mean purines and pyrimidines, $\mu^+ = \mu^-$ is a natural assumption, if they are to symbolise wild-type/mutant site, a ratio of $\mu^+/\mu^- = 3$ allows for the fact that there is only one wild-type state, but three mutant states. Here, the choice of mutation rates in the two-state model shall be restricted to the symmetric case $\mu^+ = \mu^- =: \mu$, as this is necessary to fulfil the requirement of reversibility, which is needed later, cf. equation (5.2). The mutation rates are meant to denominate rates taken over the whole sequence, such that a mutation at one particular site happens with a rate μ/N.

Whatever the choice of mutation rates, simply by focusing on single point mutations, the neighbourhood in the sequence space is defined. Each sequence has N neighbours, which match the sequence in $N - 1$ sites and differ in one. This structure in the sequence space becomes obvious by considering the *Hamming distance* $d_H(\boldsymbol{\sigma}, \boldsymbol{\sigma}')$ between

two sequences σ and σ' [Ham50, vL82]. It is defined as the number of sites at which the two sequences differ. With the single step mutation model, two sequences σ and σ' are neighbours, if and only if $d_H(\sigma, \sigma') = 1$. Consider, for instance, the sequences $\sigma_1 = 0110$, $\sigma_2 = 0100$ and $\sigma_3 = 0101$. σ_1 and σ_2 are neighbours as are σ_2 and σ_3 with $d_H(\sigma_1, \sigma_2) = d_H(\sigma_2, \sigma_3) = 1$, but $d_H(\sigma_1, \sigma_3) = 2$ and thus σ_1 and σ_3 are not neighbours.

With the definition of the Hamming distance d_H, the mutation matrix $\mathcal{M}$ reads explicitly

$$\mathcal{M}_{\sigma'\sigma} = \begin{cases} \mu/N & \text{if } d_H(\sigma', \sigma) = 1, \\ -\mu & \text{if } \sigma' = \sigma, \\ 0 & \text{otherwise.} \end{cases} \tag{3.1}$$

Again, the factor $1/N$ is chosen such that μ is the overall mutation rate, and scales intensively with the sequence length. Alternatively, one could choose μ as the per-site mutation rate, with overall mutation rate $N\mu$, leading to $-N\mu$ as diagonal entries. It is, however, important that the mutation rate and fitness function are chosen with a consistent scaling, i.e., either both intensive or extensive. For a more detailed discussion of scaling issues, see [BW01]. Note that the diagonal entries are chosen such that the mutation matrix is a Markov generator, as postulated in chapter 2.

3.2.2 Four-state models

In the case of a four letter alphabet, the situation is a little more complex. Here, there are in principle twelve different mutation rates to account for, even under the assumption of homogeneous rates across sites, because for each of the four bases there are three possible mutations. A good account of established mutation schemes is given in [SOWH96], see also [Ewe04, chapter 12]. The model that shall be used here is the *Kimura 3ST mutation scheme* [Kim81], see Figure 3.1, where there are only three different mutation rates out of the possible twelve. In particular, the forward and backward mutation rates are assumed to be identical.

The three different mutation rates correspond to *transitions* and two types of *transver-*

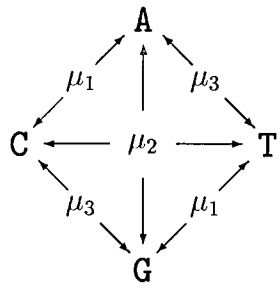

Figure 3.1: The Kimura 3ST mutation scheme.

sions. The transitions are the replacements of one purine by the other (A $\leftrightarrow$ G) or one pyrimidine by the other (C $\leftrightarrow$ T), which happen with rate μ_2. The transversions are the replacements of a purine by a pyrimidine and vice versa, which occur at rates μ_1 or μ_3, as specified in figure 3.1.

It has been observed that the mutation rate for the transitions μ_2 is considerably higher than the rates for the transversions, which in turn are rather similar $\mu_1 \approx \mu_3$. This justifies a simplification of the Kimura 3ST mutation scheme using only two different mutation rates, μ_2 and $\mu := \mu_1 = \mu_3$. The resulting scheme is known as *Kimura 2 parameter model* [Kim80].

A further specialisation is obtained by the assumption that all mutation rates coincide, so $\mu := \mu_1 = \mu_2 = \mu_3$, which is known as the *Jukes Cantor mutation scheme* [JC69].

In the four-state sequence space models, each sequence has $3N$ neighbours, as there are three different possible mutations at each site. Hence, here it is not sufficient to describe the distance between sequences $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma}'$ by the total Hamming distance $d_H(\boldsymbol{\sigma}, \boldsymbol{\sigma}')$, which just determines the number of sites at which the sequences differ, but it is necessary to observe the type of mismatch (i.e., with which mutation rate the corresponding mutation occurs). Thus a generalised three-dimensional Hamming distance $\boldsymbol{d}_H(\boldsymbol{\sigma}, \boldsymbol{\sigma}') = (d_{H,k})_{k=1,2,3}$ is required, which counts the number of mutations of each type between sequences $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma}'$. Of course the total Hamming distance is related to the generalised Hamming distance via $d_H = \sum_{k=1}^{3} d_{H,k}$.

The mutation matrix $\mathcal{M}$ for the four-state model is given by

$$\mathcal{M}_{\sigma'\sigma} = \begin{cases} \mu_1/N & \text{if } \boldsymbol{d}_H(\sigma',\sigma) = (1,0,0)^T, \\ \mu_2/N & \text{if } \boldsymbol{d}_H(\sigma',\sigma) = (0,1,0)^T, \\ \mu_3/N & \text{if } \boldsymbol{d}_H(\sigma',\sigma) = (0,0,1)^T, \\ -(\mu_1+\mu_2+\mu_3) & \text{if } \sigma' = \sigma, \\ 0 & \text{otherwise,} \end{cases} \tag{3.2}$$

where the factor $1/N$ in the single site mutation rates is used in order to obtain the μ_k as overall mutation rates, as in section 3.2.1.

3.3 Fitness functions

Whereas in the sequence space approach, mutation is straightforward to model and well justified on the molecular level, this is less so with with the selection process. This lies in the very nature of the model, as it is based on the genotype structure, which is the level at which mutation actually happens. In contrast, selection acts on the phenotype; and the mapping from genotype to phenotype is highly complex. In the alternative phenotypic approach [TLK96, BB00], the modelling is based on the phenotype rather than the genotype, but there, the mutation model is fairly arbitrary.

The phenotype that is connected to a piece of DNA depends on its function. It is useful to classify the DNA into coding and non-coding regions.

If a region codes for a protein, the genetic code applies: Every three nucleotides form a *codon* that codes for one amino acid, thus creating $4^3 = 64$ different codons for the 20 amino acids. Hence not all point-mutations in the DNA result in a change of amino acid and therefore a changed phenotype. The general picture is that the second codon position has the most impact on the amino acid and the third codon position the least, with the vast majority of transitions at the third position being silent. In addition to this, it is expected that those bits of DNA that code for the active centre of the protein have more impact on its functioning and therefore on fitness than the ones coding for the bulk part

of the protein.

For DNA regions coding for functional RNA such as t-RNA for instance, the phenotype is essentially the shape of the RNA. This shape arises by the sequence folding back onto itself and thus creating loop structures, that are mainly determined by the possible base pairing within the sequence itself. The prediction of RNA structure has been a wide area of research, and continues to be an active field, see for instance [HFS+94, FFHS00, KB03], for a review see [Hig00].

In particular, it has been shown that the mapping from genotype to structure is highly complex [FS98a, FS98b]. For instance, there are extended networks in the genotype space that fold into the same structures, in fact, any typical structure can be reached from any random structure in only a few mutational steps. On the other hand, the genotype-structure map also shows discontinuous behaviour.

The function of the so-called non-coding regions of the DNA is not entirely clear. Some parts of it certainly have some regulatory functions, the mechanisms of which are however not (or only partly) known, some parts might actually be just "junk" without any function, evolving free from selection. Due to the lack of selection, these areas of the DNA are particularly useful for determining mutation rates or for phylogenetic inference [Kon98].

Measuring the fitness experimentally is a very tricky business. To this end, *mutation accumulation experiments* have been performed since the 1960s [Muk64], where some population, for instance of the fruit fly *Drosophila*, is reproduced for a large number of generations under relaxed selection. Comparing the fitness of this population with a control population that has evolved under normal conditions yields information on the average effect of a mutation. For a review, see [Kon98].

So although there are some hints as to the general features one would expect, only little is known about the nature of realistic fitness functions. This means that the choice of the fitness function is a rather difficult business, and often driven by feasibility rather than reality.

3.3.1 Permutation-invariant fitness functions

A very common assumption, often made implicitly, is the assumption of a *permutation-invariant* fitness function. This means that the fitness of a sequence depends only on the number of mutations that it has compared to the wild-type, not on the location of the mutations, which is expressed by the Hamming distance to the wild-type. This is of course an over-simplistic assumption that by no means reflects the complexity of real fitness functions. However, in the permutation-invariant framework, the accumulation of a large number of minor deleterious mutations is described surprisingly well.

As mentioned previously, the models considered here also have an interpretation in physics, where they are mainly used to model magnetism. However, in that use, permutation-invariant interactions between sites are considered a very crude approximation, because the short range interactions, for instance next-neighbour interactions, are dominant. In contrast, in the biological setting the permutation-invariant fitness is a better approximation than a next-neighbour interaction, because here long-range interactions are incorporated, which are intrinsic in biology, since the three-dimensional structure of an RNA or protein sequence brings distant elements of the chain close together.

Permutation-invariant fitness functions come in a variety of flavours. The trivial fitness function is flat across the entire type space. As this corresponds to the case of no selection, one can tackle more complex models that also incorporate evolutionary factors other than mutation and selection; in the stochastic framework, for instance, Kimura's *neutral theory of molecular evolution* [Kim83] has been developed on the assumption of a flat fitness function to a considerable degree of sophistication.

Things get a little more complicated with a fitness function that depends linearly on the number of mutations. This is often termed *additive fitness*, because the effects of several mutations simply add up. In a discrete-time set-up, this corresponds to a multiplicative fitness. Under additive fitness, every mutation has the same effect, and there is no interaction between sites.

In principle, in the permutation-invariant setup, the fitness can be any function of the number of mutations. In any non-linear fitness function, there is some interaction between sites, called *epistasis*, albeit due to the permutation-invariance the interactions taken into account here are only average interactions. The strength of this interaction is given by the curvature of the fitness function, i.e., its second derivative. If the second derivative is negative, one speaks of *positive* or *synergistic* epistasis. Here, the effect of a mutation is stronger if there are already more mutations present in the genome. For *negative*, *antagonistic* or *diminishing returns* epistasis, i.e., a fitness function with positive second derivative, the opposite is true: a mutation that occurs on a mutant that is close to the wild-type will have a larger effect than one that occurs on a genome that already carries a lot of mutations.

Which of the two types of epistasis, synergistic or antagonistic, plays a more important role in nature is still not clear; there are arguments and studies supporting both hypotheses, see the discussion in [WA01]. It has also been argued that in the permutation-invariant set-up, the type of epistasis depends on the chosen reference sequence [GdJ93, HW01, WA01]: If this sequence lies in the centre of a cluster of high-fitness sequences, the average direction of epistasis will be synergistic, whereas choosing a reference sequence on the outskirts of such a cluster leads to antagonistic epistasis [WLA03].

The simplest example of a non-linear fitness function is quadratic (corresponding to a Gaussian fitness in discrete time), which is an example of a smooth function of the Hamming distance to the wild-type. This is different for the so-called truncation selection, a popular model in biology, see for instance [Kon88, Muk69, Cha88]. Here, the fitness is assumed to be constantly high for any number of mutations that is smaller than a critical value, and beyond that critical number it is typically assumed to be constant at a lower level. One special case of truncation selection is Eigen's *single peaked landscape* [Eig71], where the only type with high fitness is the wild-type, whereas all other types are equally unfit. It has been proposed as a model of prebiotic evolution, but even there it should only be considered as a toy model.

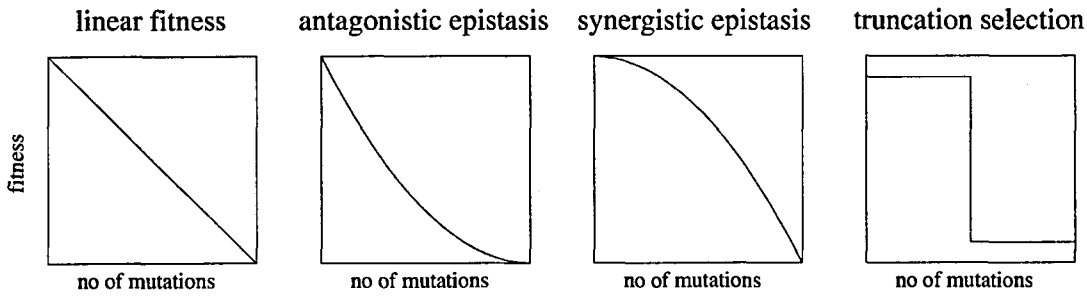

Figure 3.2: Examples of different types of permutation-invariant fitness functions.

Examples of the above mentioned types of fitness functions are visualised in Figure 3.2.

3.3.2 Spin glass models

Due to their simplicity, the permutation-invariant fitness functions have proved very popular. However, despite the possibility for using step functions, the permutation-invariant functions are still rather smooth. It is now generally believed [BG00] that due to complex interactions within and among genes, realistic fitness functions contain a certain degree of "ruggedness" that cannot be achieved by a permutation-invariant fitness function. The term ruggedness is used to describe the surface of the fitness function in sequence space, and in particular the number of saddle points and extrema as well as characteristics such as deep valleys and high ridges. Although I am not aware of a formal definition, the ruggedness of a fitness landscape has been characterised in [KL87] by considering properties such as the number of global and local optima or the expected length of an adaptive walk. How rugged the fitness landscapes in reality are is however unknown.

One possibility to incorporate ruggedness is to use the so-called spin glass models that have been used in statistical physics to model systems that contain a high degree of frustration, which means that the system is unable to find a state in which all local preferences are satisfied [MPV87]. In the context of molecular evolution, they have been suggested by Kauffman and Levin [KL87] under the name of NK landscapes, and used for instance in [CAW02]. Here, the N stands for the length of the sequence, and each spin

interacts with K other spins.

Using an alphabet $\mathcal{A} = \{+, -\}$, the NK fitness landscape has the form

$$\mathcal{R}_\sigma = \sum_{\alpha_1,\ldots,\alpha_{K+1}} J_{\alpha_1,\ldots,\alpha_{K+1}} \sigma_{\alpha_1} \cdots \sigma_{\alpha_{K+1}} \,. \tag{3.3}$$

Here, the coupling constants $J_{\alpha_1,\ldots,\alpha_{K+1}}$ are independent, identically distributed random variables. This class of fitness landscapes is tunably rugged through K, i.e., by varying the number of spins K that interact with any given spin, the ruggedness can be tuned. The larger K, the more rugged the fitness landscape is. For $K = 1$, one has the Sherrington-Kirkpatrick spin glass [SK75], see also [Tal03], which has first been suggested in the evolutionary context by Anderson [And83].

A special case of the Sherrington-Kirkpatrick spin glass is the Hopfield Hamiltonian [Hop82], which has first been suggested as a model for neural networks. It can be obtained by choosing the couplings as

$$J_{\alpha\beta} = \sum_{q=0}^{p} \xi_\alpha^q \xi_\beta^q \,, \tag{3.4}$$

where the $\boldsymbol{\xi}^q$, with $q = 0, \ldots, p$ are predefined sequences, or *patterns*, and the fitness of a particular sequence depends on its overlap with these patterns. Note that in the set-up as described in equation (3.3) with (3.4), the global maxima of the fitness landscape coincide with the patterns and their complementary sequences.

Usually, the patterns are chosen randomly. However, once they are chosen, the system is kept fixed at this choice, which is sometimes called the case of "quenched" randomness. The statistical properties of this landscape have been studied in detail in [AGS85a, AGS85b], see also [Tal03]. In the thermodynamic limit $N \to \infty$ with finitely many patterns, there are $2(p+1)$ global maxima, perfectly correlated with the patterns $\{\xi^q\}$, and their complements $\{-\xi^q\}$. In addition to that, the number of local maxima and saddle points grows exponentially with the number of the patterns $p+1$. Hence the ruggedness of the Hopfield-fitness can be tuned by varying the number of patterns.

Using only one pattern, one is again in a permutation-invariant set-up with a quadratic fitness, where the wild-type is given by the pattern, as in this case the fitness can be written

as

$$\mathcal{R}_{\sigma} = \sum_{\alpha,\beta} \xi_{\alpha}\xi_{\beta}\sigma_{\alpha}\sigma_{\beta} = \left(\sum_{\alpha} \xi_{\alpha}\sigma_{\alpha}\right)^{2} . \tag{3.5}$$

This thesis investigates permutation-invariant fitness function alongside generalised Hopfield-type fitness functions, which are introduced in section 4.2. The reason for investigating the Hopfield-type fitness is that it can be treated within the framework set out here, but it does display a certain degree of ruggedness that is thought to be important in biologically realistic fitness functions.

CHAPTER 4

Lumping: Reduction of the sequence space

One problem of the sequence space approach is the large number of types, which grows exponentially with the sequence length N, $|\mathfrak{S}| = |\mathcal{A}|^N$. The time-evolution operator $\mathcal{H}$ is a matrix of size $|\mathfrak{S}| \times |\mathfrak{S}|$, and in this set-up one is interested in its leading eigenvalue $\bar{r}$ and the corresponding right and left eigenvectors $\boldsymbol{p}$ and $\boldsymbol{z}$.

The relevant sequence length depends on the particular application one has in mind, but it is typically rather long. If one aims to model the whole genome of a virus or a bacterium, N has to be in the region of $N \approx 10^6$, but even a single gene has of the order of $N \approx 10^3$ base pairs. These values lead to matrices of a size that makes the eigenvalues and eigenvectors inaccessible.

For some types of fitness functions, this problem can be reduced by *lumping* together types into *classes* of types, and considering the new process on a reduced sequence space, which contains the classes rather than the individual types. Under certain circumstances, mutation is described as a Markov process in the emerging lumped process as well, such that this process is accessible to Markov process methods, and the theory developed in chapter 2 can also be applied to the lumped system.

The lumping of the mutation process is a standard procedure in the theory of Markov chains, [KS60, chapter 6], see also [BBBK05] for an application to mutation–selection models. This lumping leads to a meaningful mutation–selection model on the reduced type space, if all sequences lumped together into one class have the same fitness.

It is possible to lump the Markov chain given by the mutation matrix $\mathcal{M}$ with state space $\mathfrak{S}$ with respect to a particular partition $\mathfrak{S} = \dot{\bigcup}_{k=0}^{r} \mathfrak{S}_k$, if and only if for each pair $\mathfrak{S}_k, \mathfrak{S}_\ell$ the cumulative mutation rates $u_{\mathfrak{S}_\ell, i} = \sum_{j \in \mathfrak{S}_\ell} \mathcal{M}_{ji}$ from type $i \in \mathfrak{S}_k$ into $\mathfrak{S}_\ell$, are

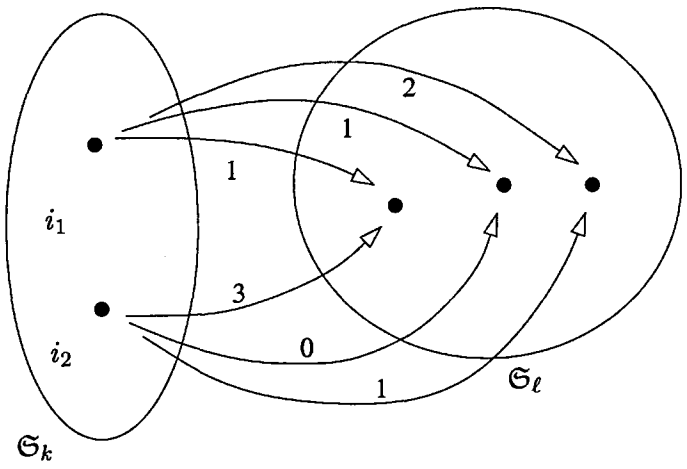

Figure 4.1: Visualisation of the compatibility with lumping: Consider two classes $\mathfrak{S}_k$ and $\mathfrak{S}_\ell$. The mutation rates from the types in $\mathfrak{S}_k$ to the types in $\mathfrak{S}_\ell$ (given next to the arrows) are compatible with a lumping with respect to $\mathfrak{S}_k$ and $\mathfrak{S}_\ell$, because the sum of the mutation rates from type i_1 to all types in $\mathfrak{S}_\ell$ is given by $u_{\mathfrak{S}_\ell,i_1} = 1+1+2 = 4$, which is identical with those from i_2, $u_{\mathfrak{S}_\ell,i_2} = 3+0+1 = 4$.

identical for all $i \in \mathfrak{S}_k$, cf. the example shown in figure 4.1. In this case, the lumped process, with $r+1$ states $\mathfrak{S}_0, \ldots, \mathfrak{S}_r$ and mutation rates $u_{\mathfrak{S}_k,\mathfrak{S}_\ell}$, is again a Markov chain [KS60, theorem 6.3.2].

In what follows, the lumping procedure shall be performed for permutation-invariant and Hopfield-type fitness functions, both for two- and four-state models, using a consistent notation. A two-state model with permutation-invariant fitness has been considered in [HRWB02], as an effectively lumped system, but since in this case the lumping is fairly simple, it has been done implicitly. Here, it shall be included as a simple example merely to demonstrate the method, which is essential for the more complex cases. The lumping for a four-state model with permutation-invariant fitness has been done in [GG04b].

Lumping for a two-state model with Hopfield fitness has been performed in [BABG03, BBBK05], and this shall be applied here to a generalised *Hopfield-type fitness*. A model with more than two states has been used in [Gay92], but this has a different symmetry than the one induced by the Kimura 3ST mutation scheme. The approach taken here goes in line with the permutation-invariant case performed in this work and differs from the method used in [Gay92].

4.1 Permutation-invariant fitness

4.1.1 The two-state model with permutation-invariant fitness

The alphabet $\mathcal{A} = \{0, 1\}$ shall be used throughout for the two-state model. The mutation matrix is given in equation (3.1). An arbitrary sequence is chosen as reference sequence $\boldsymbol{\sigma}_{\text{ref}}$, and represented by the string $00\ldots0$. Usually (but not necessarily) the reference sequence coincides with the *wild-type*, i.e., the sequence with maximal fitness. Here, the terms reference sequence and wild-type shall be used synonymously.

For each sequence $\boldsymbol{\sigma}$, the Hamming distance to the wild-type is of particular interest and termed the *mutational distance* $d_{\boldsymbol{\sigma}}$ to the wild-type, with $d_{\boldsymbol{\sigma}} := d_H(\boldsymbol{\sigma}_{\text{ref}}, \boldsymbol{\sigma}) = \sum_{\alpha=1}^{N} \sigma_\alpha$. The mutational distance takes values $0, \ldots, N$. As the permutation-invariant fitness does not depend on the order of the sequence, but only on the number of mutations, it is a function of the mutational distance only, $\mathcal{R}_{\boldsymbol{\sigma}} = \mathcal{R}(d_{\boldsymbol{\sigma}})$. This suggests a lumping with respect to the mutational distance.

The relevant partition of the sequence space is given by $\mathfrak{S} = \dot{\bigcup}_{d=0}^{N} \mathfrak{S}_d$ with $\mathfrak{S}_d = \{\boldsymbol{\sigma} | \, d_{\boldsymbol{\sigma}} = d\}$, and the reduced sequence space, or *mutational distance space* $\mathcal{S}$, contains the classes $\mathcal{S} = \{0, 1, \ldots, N\}$. The number of sequences n_d that are lumped into class d is given by the binomial coefficients,

$$n_d = \binom{N}{d} = \frac{N!}{d!(N-d)!} \,. \tag{4.1}$$

Table 4.1 shows as an example all sequences of length $N = 4$, which are lumped into the classes specified by the mutational distance d.

In the single step mutation model, the only neighbours of a sequence $\boldsymbol{\sigma}$ with mutational distance d lie in the classes $d \pm 1$. Thus the only non-zero cumulative mutation rates are

$$u_{d\pm1,\boldsymbol{\sigma}} = u_d^{\pm} = \sum_{\boldsymbol{\sigma}' \in \mathfrak{S}_{d\pm1}} \mathcal{M}_{\boldsymbol{\sigma}'\boldsymbol{\sigma}} \,. \tag{4.2}$$

Consider, for instance, the sequence $\boldsymbol{\sigma} = 1000 \in \mathfrak{S}_1$. This sequence has three neighbours in $\mathfrak{S}_2$, as indicated by the arrows in table 4.1, because a mutation at each of its 0-sites creates a different sequence with mutational distance $d = 2$. Thus the cumulative mutation

d	0	1	2	3	4
σ	0000	1000 0100 0010 0001	1100 1010 1001 0110 0101 0011	1110 1101 1011 0111	1111
n_d	1	4	6	4	1

Table 4.1: Enumeration of all two-state sequences of length $N = 4$ and their lumping into mutational distance classes d. The arrows point to the neighbours of two selected sequences within one neighbouring class.

rate from sequence σ into $\mathfrak{S}_2$ is given by $u_1^+ = 3\mu/4$ (as the single site mutation rates $\mathcal{M}_{\sigma'\sigma} = \mu/N$ for neighbouring sequences). This is in fact true for any sequence in $\mathfrak{S}_1$, fulfilling the criterion for lumping; the only differences between the sequences within $\mathfrak{S}_1$ are the particular sequences in $\mathfrak{S}_2$, which are their neighbours. Similarly, a sequence in $\mathfrak{S}_2$, say $\sigma = 0101$, has two neighbours in $\mathfrak{S}_1$, as it has two 1-sites at which it can mutate to become a sequence in $\mathfrak{S}_1$, which is the case for any sequence in $\mathfrak{S}_2$. Hence, the backward cumulative mutation rate from $\mathfrak{S}_2$ to $\mathfrak{S}_1$ is given by $u_2^- = 2\mu/4$.

More generally, any fixed sequence σ in $\mathfrak{S}_d$ has d entries 1, so there are d different sequences σ' in $\mathfrak{S}_{d-1}$ such that $d_H(\sigma, \sigma') = 1$ and thus $\mathcal{M}_{\sigma'\sigma} = \mu/N$. Hence the cumulative mutation rate from any $\sigma \in \mathfrak{S}_d$ into $\mathfrak{S}_{d\pm1}$ is, independently of the order of the sequence σ, given by

$$\begin{aligned} u_d^- &:= u_{d-1,\sigma} = \mu\, d/N\,, \\ u_d^+ &:= u_{d+1,\sigma} = \mu(N-d)/N\,, \end{aligned} \tag{4.3}$$

where u_d^+ is obtained by an equivalent argument for $d \to d+1$. As these cumulative mutation rates are the same for any sequence $\sigma \in \mathfrak{S}_d$, independently of its order, the Markov chain describing the mutation matrix is lumpable with respect to the partition of $\mathfrak{S}$ induced by the permutation-invariant fitness, and the number of types can be reduced from $|\mathfrak{S}| = 2^N$ to $|\mathcal{S}| = N + 1$.

The mutation–selection process on the mutational distance space $\mathcal{S}$ is described by the time-evolution operator $\boldsymbol{H} = \boldsymbol{R} + \boldsymbol{M}$, with lumped reproduction and mutation matrices

$\boldsymbol{R}$ and $\boldsymbol{M}$ of dimension $|\mathcal{S}| \times |\mathcal{S}|$. These can be obtained from the original matrices $\mathcal{R}$ and $\mathcal{M}$ by the means of a transformation $\boldsymbol{L}$, which is an $|\mathfrak{S}| \times |\mathcal{S}|$ dimensional matrix that specifies how the sequences are lumped into the mutational distance classes. To this end, label the columns of $\boldsymbol{L}$ with the mutational distance d, the rows with the actual sequence σ. The entries of $\boldsymbol{L}$ are given by

$$L_{\sigma d} = \begin{cases} 1/n_d & \text{if } d_\sigma = d, \\ 0 & \text{otherwise.} \end{cases} \tag{4.4}$$

Thus $\sum_\sigma L_{\sigma d} = 1$, which ensures that the Markov property is preserved. This is not the only possible choice of $\boldsymbol{L}$; an alternative would be to have entries 1 in each column only in the first row where $d_\sigma = d$, but it seems more natural to superimpose all sequences that contribute to a class as done above.

Although only square matrices are invertible, by abuse of notation, the matrix $\boldsymbol{L}^{-1}$ shall denote an $|\mathcal{S}| \times |\mathfrak{S}|$-dimensional matrix such that $\boldsymbol{L}^{-1}\boldsymbol{L} = \mathbf{1}_{|\mathcal{S}|}$ and $\boldsymbol{L}\boldsymbol{L}^{-1}$ is block-diagonal, where each block corresponds to all sequences that share the same mutational distance, and $\sum_\sigma (\boldsymbol{L}\boldsymbol{L}^{-1})_{\sigma\sigma'} = \sum_{\sigma'} (\boldsymbol{L}\boldsymbol{L}^{-1})_{\sigma\sigma'} = 1$. This can be fulfilled by

$$L^{-1}_{d\sigma} = \begin{cases} 1 & \text{if } d_\sigma = d, \\ 0 & \text{otherwise.} \end{cases} \tag{4.5}$$

Here again, the choice of $\boldsymbol{L}^{-1}$ is not unique, even for fixed $\boldsymbol{L}$, but this choice is convenient.

With this, the reduced time-evolution operator is given by

$$\boldsymbol{H} = \boldsymbol{R} + \boldsymbol{M} = \boldsymbol{L}^{-1}(\mathcal{R} + \mathcal{M})\boldsymbol{L}\,. \tag{4.6}$$

Performing this transformation, of course the off-diagonal entries of $\boldsymbol{M}$ match the cumulative mutation rates given in equation (4.3), whereas the diagonal entries of both $\boldsymbol{R}$ and $\boldsymbol{M}$ are unchanged in the sense that $(\boldsymbol{R} + \boldsymbol{M})_{d_\sigma d_\sigma} = (\mathcal{R} + \mathcal{M})_{\sigma\sigma}$ for any sequence σ with

$$\begin{aligned} \boldsymbol{\sigma}_{\text{ref}} &= \ldots\texttt{AGCCATA}\ldots \;\rightarrow\; \ldots\texttt{0000000}\ldots \\ \boldsymbol{\sigma} &= \ldots\texttt{AACGATA}\ldots \;\rightarrow\; \ldots\texttt{0203000}\ldots \end{aligned}$$

Figure 4.2: Sequence $\boldsymbol{\sigma}$ is mapped from the natural alphabet $\{\texttt{A},\texttt{G},\texttt{C},\texttt{T}\}$ to $\mathcal{A} = \{0,1,2,3\}$ by comparing it to a given reference sequence $\boldsymbol{\sigma}_{\text{ref}}$ and assigning 0 at sites where they match and $1,2,3$ at sites where they differ, depending on the type of mutation given by the Kimura 3ST mutation scheme from figure 3.1.

mutational distance $d_{\boldsymbol{\sigma}}$, and thus

$$M_{d'd} = \begin{cases} u_d^+ & \text{if } d' = d+1, \\ u_d^- & \text{if } d' = d-1, \\ -\mu & \text{if } d' = d, \\ 0 & \text{otherwise.} \end{cases} \tag{4.7}$$

Note that $\boldsymbol{H}$ is a time-evolution operator acting on $\mathcal{S}$, and describes the evolution of a population under mutation and selection determined by the evolution equation (2.1), and thus the theory developed in chapter 2 applies.

4.1.2 The four-state model with permutation-invariant fitness

In the four-state model, it is convenient to deviate from the natural alphabet $\{\texttt{A},\texttt{G},\texttt{C},\texttt{T}\}$ and use $\mathcal{A} = \{0,1,2,3\}$ with $\boldsymbol{\sigma}_{\text{ref}} = 00\ldots0$ chosen as the wild-type. The mapping of any other sequence $\boldsymbol{\sigma}$ from the natural alphabet into $\mathcal{A}$ can be obtained by assigning 0 at each site where it matches the reference sequence $\boldsymbol{\sigma}_{\text{ref}}$, and assigning 1, 2, or 3 at any other site, depending on the type of mutation by which it differs. For an example, see figure 4.2.

Because in the four-state model a generalised three-dimensional Hamming distance is used, it is also necessary to use a vectorial mutational distance $\boldsymbol{d}$ of sequence $\boldsymbol{\sigma}$ to the wild-type, $\boldsymbol{d} = (d_k)_{k=1,2,3} \in \mathbb{Z}^3$ with

$$d_k := d_{H,k}(\boldsymbol{\sigma}_{\text{ref}}, \boldsymbol{\sigma}) = \sum_{\alpha=1}^{N} \delta(k, \sigma_\alpha) \quad \text{with } k \in \{1,2,3\}, \tag{4.8}$$

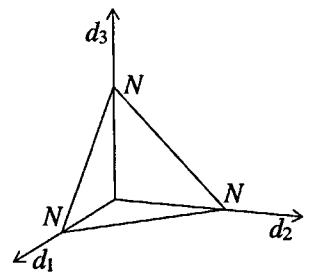

Figure 4.3: Visualisation of the structure of the mutational distance space $\mathcal{S}$.

using the Kronecker symbol $\delta(.,.)$. One has

$$0 \leq d_k \quad \text{and} \quad \sum_{k=1}^{3} d_k \leq N\,, \tag{4.9}$$

and thus there are $|\mathcal{S}| = (N+1)(N+2)(N+3)/6$ different mutational distances. The number of wild-type sites d_0 is given by $d_0 = \sum_{\alpha=1}^{N} \delta(0,\sigma_\alpha) = N - \sum_{k=1}^{3} d_k$.

As in the two-state case, the permutation-invariant fitness is a function of the mutational distance $\boldsymbol{d}$ only, and again, the lumping shall be performed with respect to $\boldsymbol{d}$. The mutational distance space $\mathcal{S}$ contains all mutational distances $\boldsymbol{d} \in \mathbb{Z}^3$ that fulfil (4.9). Thus $\mathcal{S}$ is a simplex in $\mathbb{Z}^3$, cf. figure 4.3. The partition induced on the original sequence space is $\mathfrak{S} = \dot{\bigcup}_{\boldsymbol{d}\in\mathcal{S}} \mathfrak{S}_{\boldsymbol{d}}$ with $\mathfrak{S}_{\boldsymbol{d}} = \{\boldsymbol{\sigma} | \boldsymbol{d}_{\boldsymbol{\sigma}} = \boldsymbol{d}\}$. The number of sequences $n_{\boldsymbol{d}}$ in each mutational distance class $\boldsymbol{d}$ is given by the multinomial coefficients

$$n_{\boldsymbol{d}} = \binom{N}{d_0, d_1, d_2, d_3} = \frac{N!}{d_0! d_1! d_2! d_3!}\,, \tag{4.10}$$

where $d_0 + d_1 + d_2 + d_3 = N$.

In this case, the structure of the neighbourhood is more complicated than in the two-state model. A sequence with mutational distance $\boldsymbol{d}$ has all its neighbours in the classes $\boldsymbol{d} \pm \boldsymbol{e}_k$ and $\boldsymbol{d} \pm \boldsymbol{e}_k \mp \boldsymbol{e}_\ell$, $k < \ell \in \{1,2,3\}$, where the $\boldsymbol{e}_k$ and $\boldsymbol{e}_\ell$ are the Cartesian unit vectors $\boldsymbol{e}_1 = (1,0,0)^T$, $\boldsymbol{e}_2 = (0,1,0)^T$ and $\boldsymbol{e}_3 = (0,0,1)^T$, because these are the only mutations that can be obtained by changing a single site. The $\boldsymbol{e}_k$ and $\boldsymbol{e}_k - \boldsymbol{e}_\ell$ (with $k < \ell$) shall collectively be called the *unit vectors of mutation* $\boldsymbol{e}_\chi$. As there are six unit vectors of mutation, each of which can be applied in forward and backward direction, each mutational distance $\boldsymbol{d}$ has twelve neighbouring mutational distances $\boldsymbol{d} \pm \boldsymbol{e}_\chi$ (except those

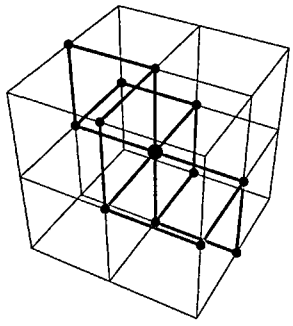

Figure 4.4: Sketch of the possible mutational directions in the permutation-invariant four-state model. The blue lines show the cubic lattice of the mutational distance space $\mathcal{S}$, the central dot is the mutational distance $\boldsymbol{d}$ under consideration, and the smaller dots are its neighbours.

on the borders of the mutational distance space $\mathcal{S}$), cf. the sketch in figure 4.4.

The cumulative mutation rate from a sequence $\boldsymbol{\sigma}$ with $d_{\boldsymbol{\sigma}}$ into class $\boldsymbol{d} \pm \boldsymbol{e}_\chi$ is indeed independent of the particular sequence: For instance a mutation from $\boldsymbol{d}$ to $\boldsymbol{d} + \boldsymbol{e}_k$ is a mutation that increases d_k by one and thus is the mutation of a wild-type site into type k, which happens with rate μ_k/N. As there are d_0 different wild-type sites in the sequence $\boldsymbol{\sigma}$, all of which can mutate with the desired effect, the cumulative mutation rate is given by $u_{\boldsymbol{d}}^{+k} = \mu_k d_0/N$, independently of the order of the sequence. Similarly, we get for all cumulative mutation rates $u_{\boldsymbol{d}}^{\pm\chi}$

$$\begin{aligned} u_{\boldsymbol{d}}^{+k} &= \mu_k d_0/N && \text{for } \boldsymbol{d} \to \boldsymbol{d} + \boldsymbol{e}_k, \\ u_{\boldsymbol{d}}^{-k} &= \mu_k d_k/N && \text{for } \boldsymbol{d} \to \boldsymbol{d} - \boldsymbol{e}_k, \\ u_{\boldsymbol{d}}^{+k,-\ell} &= \mu_m d_\ell/N && \text{for } \boldsymbol{d} \to \boldsymbol{d} + \boldsymbol{e}_k - \boldsymbol{e}_\ell, \end{aligned} \tag{4.11}$$

with $\{k, \ell, m\} = \{1, 2, 3\}$. This notation is a short way to denote that the k, ℓ, m are distinct numbers in $\{1, 2, 3\}$. Note that the mutation rate from a mutational distance at the border of the mutational distance space $\mathcal{S}$ to the outside vanishes with equation (4.11), as required. Hence, the evolution process is lumpable with respect to the partition induced by the mutational distances, i.e., the lumped process is again Markovian, and the lumped mutation rates are given in equation (4.11).

The lumped time-evolution operator $\boldsymbol{H}$ can be obtained in the same way as in the two-state model, with equations (4.4), (4.5) and (4.6) applying in the four-state case if

three-dimensional mutational distances $\boldsymbol{d}$ and the $n_{\boldsymbol{d}}$ from equation (4.10) are used. This leads to a lumped mutation matrix M given as

$$M_{\boldsymbol{d}'\boldsymbol{d}} = \begin{cases} u_{\boldsymbol{d}}^{+\chi} & \text{if } \boldsymbol{d}' = \boldsymbol{d} + \boldsymbol{e}_\chi, \\ u_{\boldsymbol{d}}^{-\chi} & \text{if } \boldsymbol{d}' = \boldsymbol{d} - \boldsymbol{e}_\chi, \\ -\sum_\chi (u_{\boldsymbol{d}}^{+\chi} + u_{\boldsymbol{d}}^{-\chi}) & \text{if } \boldsymbol{d}' = \boldsymbol{d}, \\ 0 & \text{otherwise,} \end{cases} \tag{4.12}$$

with the $u_{\boldsymbol{d}}^{\pm\chi}$ from equation (4.11).

4.2 Hopfield-type fitness

The Hopfield Hamiltonian as described in section 3.3.2 was first introduced in [Hop82] as a model for neural networks. It is a special case of the Sherrington-Kirkpatrick spin glass, which was first suggested in the context of molecular evolution by [And83]. In the works of [Leu87, Tar92], the Hopfield Hamiltonian was used as an example of a non-permutation-invariant fitness in a discrete-time formulation of a mutation–selection model in the sequence space approach.

Formulated for a two-state model for sequences $\boldsymbol{\sigma}$ of length N written in an alphabet $\mathcal{A} = \{+1, -1\}$, the Hopfield Hamiltonian in its usual form reads

$$H_N(\boldsymbol{\sigma}) = \frac{1}{2N} \sum_{q=0}^{p} \sum_{\alpha,\beta=1}^{N} \xi_\alpha^q \xi_\beta^q \sigma_\alpha \sigma_\beta \,, \tag{4.13}$$

where the $\boldsymbol{\xi}^q$ are $p+1$ (usually) randomly chosen, but fixed sequences that serve as predefined patterns, similarly to the single wild-type in the permutation-invariant case. In fact, in the case of a single pattern, $p = 0$, one is in the permutation-invariant setting with the pattern as wild-type.

If it is used in the evolutionary context, the Hopfield Hamiltonian serves as fitness function, $\mathcal{R}_{\boldsymbol{\sigma}} = H_N(\boldsymbol{\sigma})$. Writing equation (4.13) as

$$H_N(\boldsymbol{\sigma}) = \frac{1}{2N} \sum_{q=0}^{p} \left(\sum_{\alpha=1}^{N} \xi_\alpha^q \sigma_\alpha \right)^2 = \frac{1}{2N} \sum_{q=0}^{p} (o^q(\boldsymbol{\sigma}))^2 \tag{4.14}$$

with the overlaps $o^q(\sigma)$ of sequence σ with pattern ξ^q given as scalar product

$$o^q(\sigma) := \xi^q \cdot \sigma = \sum_{\alpha=1}^{N} \xi^q_\alpha \sigma_\alpha \,, \tag{4.15}$$

it is clear that the original Hopfield Hamiltonian is a quadratic function of these overlaps with the chosen patterns.

It is not important, which particular alphabet is used, as one can transform from one notation to the other via a linear transformation. In order to have a consistent notation, here we shall deviate from the alphabet $\{+,-\}$, which is used in wide part of the literature on spin glass models, and use alphabets $\mathcal{A} = \{0,1\}$ and $\mathcal{A} = \{0,1,2,3\}$ for the two- and four-state case, respectively. Instead of the overlaps, the model shall be formulated in terms of the *specific distances* w^q, which are the Hamming distances with respect to pattern ξ^q,

$$w^q(\sigma) := d_H(\xi^q, \sigma) \,. \tag{4.16}$$

For two-state models, the w^q are scalars, whereas for four-state models, they have three components like the mutational distances for the four-state set-up as in section 4.1.2.

As the lumping procedure for the Hopfield fitness does not depend on the quadratic nature of the Hopfield Hamiltonian, it is possible to generalise the approach to a *Hopfield-type fitness function*, which is any function that depends on the sequence only through the specific distances w^q, i.e., $\mathcal{R}_\sigma = \mathcal{R}(w^q(\sigma))$, an approach that has also been taken in [Pel02], using a truncation selection of Hopfield-type.

4.2.1 The two-state model with Hopfield-type fitness

For the Hopfield-type fitness, the same alphabet $\mathcal{A} = \{0,1\}$ and mutation models (3.1) as in the permutation-invariant set-up are used. One major difference is, however, that with a permutation-invariant fitness function, there is one reference sequence chosen and mapped onto the string $00\ldots0$, whereas with a Hopfield-type fitness one has to deal with $p+1$ patterns. However, without loss of generality, one pattern, say ξ^0, can always be chosen as $\xi^0 = 00\ldots0$. As an example, consider a case with sequence length $N = 12$ and

three patterns, $p = 2$, where the patterns ξ^1 and ξ^2 are given by $\xi^1 = (101101011100)$ and $\xi^2 = (100011111110)$.

In order to express the specific distances w^q in terms of the mutational distance to pattern ξ^0, consider the $(p+1) \times N$ matrix ξ, which collects the patterns such that the qth row of ξ is pattern ξ^q. The matrix ξ corresponding to the example given above is

$$\xi = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} \xi^0 \\ \xi^1 \\ \xi^2 \end{pmatrix} . \tag{4.17}$$

Considering the columns ξ_α of ξ, there are only 2^p different possible columns, because ξ^0 is chosen as $00\ldots0$ and thus the top entry is always 0. These distinct column vectors are collected in a $(p+1) \times 2^p$ matrix ρ. In the case $p = 2$, this matrix is given by

$$\rho = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix} \tag{4.18}$$

Using the column vectors ρ_v of the matrix ρ, the patterns given in equation (4.17) can alternatively be expressed as

$$\xi = (\rho_4, \rho_1, \rho_3, \rho_3, \rho_2, \rho_4, \rho_2, \rho_4, \rho_4, \rho_4, \rho_2, \rho_1) \ , \tag{4.19}$$

classifying the sites into 2^p classes according to which of the column vectors ρ_v of ρ coincides with the column vector ξ_α of the patterns ξ at site α. Let $\Lambda = \{1, \ldots, N\}$ be the index set of sites, with a partition into 2^p subsets Λ_v induced by the patterns ξ such that

$$\Lambda = \dot{\bigcup}_{v=1}^{2^p} \Lambda_v \, , \tag{4.20}$$

with

$$\alpha \in \Lambda_v \iff \xi_\alpha = \rho_v \, . \tag{4.21}$$

Patterns can be characterised by the number of sites $N_v = |\Lambda_v|$ in each subset.

The example patterns (4.17) can therefore be described by

$$\Lambda_1 = \{2, 12\}, \qquad \Lambda_2 = \{5, 7, 11\}, \qquad \Lambda_3 = \{3, 4\}, \qquad \Lambda_4 = \{1, 6, 8, 9, 10\},$$
$$N_1 = 2, \qquad N_2 = 3, \qquad N_3 = 2, \qquad N_4 = 5. \tag{4.22}$$

Defining the 2^p *partial distances* d_v, where $v = 1, \ldots, 2^p$, with respect to pattern $\boldsymbol{\xi}^0$ as

$$d_v := \sum_{\alpha \in \Lambda_v} \delta(1, \sigma_\alpha) \quad \text{with} \quad 0 \leq d_v \leq N_v, \tag{4.23}$$

the partial distances d_v^q with respect to any other pattern $\boldsymbol{\xi}^q$ are given by

$$d_v^q = \sum_{\alpha \in \Lambda_v} \delta(1 - \rho_v^q, \sigma_\alpha) = \begin{cases} d_v & \text{if } \rho_v^q = 0, \\ N_v - d_v & \text{if } \rho_v^q = 1, \end{cases} \tag{4.24}$$

using the matrix elements ρ_v^q of $\boldsymbol{\rho}$. The specific distance w^q to any pattern $\boldsymbol{\xi}^q$ can be expressed as

$$w^q = \sum_{v=1}^{2^p} \sum_{\alpha \in \Lambda_v} \delta(1 - \rho_v^q, \sigma_\alpha) = \sum_{v \in V_0^q} d_v + \sum_{v \in V_1^q} (N_v - d_v), \tag{4.25}$$

where the index set $V = \{1, \ldots, 2^p\}$ of the vectors $\boldsymbol{e}_v$ is partitioned into two subsets, $V = V_0^1 \cup V_1^q$ with $V_0^q = \{v | \rho_v^q = 0\}$ and $V_1^q = \{v | \rho_v^q = 1\}$.

For any arbitrary sequence of length $N = 12$, say

$$\boldsymbol{\sigma} = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}, \tag{4.26}$$

the partial distances d_v with respect to pattern $\boldsymbol{\xi}^0$ are just given by the number of 1's in the respective partition Λ_v, such that for the partition given by (4.22), they are given by

$$d_1 = 1, \quad d_2 = 1, \quad d_3 = 2, \quad d_4 = 3, \tag{4.27}$$

whereas the partial distances to pattern $\boldsymbol{\xi}^1$ are given by

$$d_1^1 = d_1 = 1, \quad d_2^1 = d_2 = 1, \quad d_3^1 = N_3 - d_3 = 0, \quad d_4^1 = N_4 - d_4 = 2, \tag{4.28}$$

because in Λ_1 and Λ_2, $\boldsymbol{\xi}^1$ coincides with $\boldsymbol{\xi}^0$, whereas in Λ_3 and Λ_4, $\boldsymbol{\xi}^1$ and $\boldsymbol{\xi}^0$ are complementary to each other, i.e., where one pattern has entry 0, the other has entry 1. For pattern $\boldsymbol{\xi}^1$, the index set of 0-sites is $V_0^1 = \{\Lambda_1, \Lambda_2\}$ and the index set of 1-sites is

$V_1^1 = \{\Lambda_3, \Lambda_4\}$ such that the specific distance w^1 of sequence $\boldsymbol{\sigma}$ from equation (4.26) to pattern $\boldsymbol{\xi}^1$ from equation (4.17) is given by $w^1 = d_1 + d_2 + (N_3 - d_3) + (N_4 - d_4) = 4$.

Hence, by specifying the 2^p partial distances d_v with respect to pattern $\boldsymbol{\xi}^0$, the specific distances w^q with respect to any pattern $\boldsymbol{\xi}^q$ are determined, which in turn determine the fitness. Thus the d_v, collected in a distance vector $\boldsymbol{d} = (d_v)_{v=1\ldots 2^p}$, shall be the quantities that label the sequences in the lumped system. To check whether this lumping is compatible with the mutation model, consider a sequence with distance $\boldsymbol{d}$. The only sequences that can be reached with a single mutation are those with $\boldsymbol{d} \pm \boldsymbol{e}_v$, where the $\boldsymbol{e}_v = (\delta_{vw})_{w=1\ldots 2^p}$ are the unit vectors of mutation. As a sequence with $\boldsymbol{d}$ has $(N_v - d_v)$ 0-sites and d_v 1-sites in Λ_v, the cumulative mutation rates are given by

$$\begin{aligned} u_{\boldsymbol{d}}^{+v} &= \mu\,(N_v - d_v)/N && \text{for } \boldsymbol{d} \to \boldsymbol{d} + \boldsymbol{e}_v \text{ and} \\ u_{\boldsymbol{d}}^{-v} &= \mu\, d_v/N && \text{for } \boldsymbol{d} \to \boldsymbol{d} - \boldsymbol{e}_v, \end{aligned} \tag{4.29}$$

irrespective of the particular order within the subsequences $\boldsymbol{\sigma}_v = (\sigma_\alpha)_{\alpha\in\Lambda_v}$. Therefore the cumulative mutation rates $u_{\boldsymbol{d}}^{\pm v}$ are the same for all sequences $\boldsymbol{\sigma}$ with the same $\boldsymbol{d}$, which is the condition for lumping.

Considering the sequence $\boldsymbol{\sigma}$ given in equation (4.26), and, for instance, a mutation in Λ_4, the cumulative mutation rates are

$$u_{\boldsymbol{d}}^{+4} = \mu(5-3)/N = \mu/6 \quad \text{and} \quad u_{\boldsymbol{d}}^{-4} = \mu\,3/N = \mu/4, \tag{4.30}$$

because of the five sites in Λ_4, $\boldsymbol{\sigma}$ has three entries 1, which, if they mutate, lead to a sequence with $\boldsymbol{d} - \boldsymbol{e}_4$, and thus the corresponding mutation rate is $u_{\boldsymbol{d}}^{-4}$, whereas a mutation at one of the remaining two 0-entries leads to a mutation to $\boldsymbol{d} + \boldsymbol{e}_4$, which happens at rate $u_{\boldsymbol{d}}^{+4}$. The same result is obtained for any other sequence with the same partial distances d_v given in equation (4.27).

To establish the transformation $\boldsymbol{L}$, that lumps the original sequence space $\mathfrak{S}$ onto the lumped sequence space $\mathcal{S}$, the number of different mutational distances $\boldsymbol{d}$, and the number of different sequences $\boldsymbol{\sigma}$ that are lumped onto each $\boldsymbol{d}$ are required.

To this end, consider the subsequences $\boldsymbol{\sigma}_v$: There are $N_v + 1$ possible different d_v (as d_v takes values from 0 to N_v), and there are $\binom{N_v}{d_v}$ sequences $\boldsymbol{\sigma}_v$ for each d_v. Hence, considering all sites, there are

$$|\mathcal{S}| = \prod_{v=1}^{2^p} (N_v + 1) \tag{4.31}$$

different $\boldsymbol{d}$, and

$$n_{\boldsymbol{d}} = \prod_{v=1}^{2^p} \binom{N_v}{d_v} \tag{4.32}$$

sequences $\boldsymbol{\sigma}$ that are mapped onto each $\boldsymbol{d}$. For the patterns $\boldsymbol{\xi}$ chosen as example (4.17), we have $|\mathcal{S}| = 3 \cdot 4 \cdot 3 \cdot 6 = 216$, while the full sequence space has dimension $|\mathfrak{S}| = 2^{12} = 4096$. The mutation rates can be calculated from equation (4.29).

With this, the transformations $\boldsymbol{L}$ and $\boldsymbol{L}^{-1}$ and the lumped time evolution operator $\boldsymbol{H}$ are given as in equations (4.4), (4.5) and (4.6), respectively, using the $n_{\boldsymbol{d}}$ from equation (4.32), with a lumped mutation matrix as

$$M_{\boldsymbol{d}'\boldsymbol{d}} = \begin{cases} u_{\boldsymbol{d}}^{+v} & \text{if } \boldsymbol{d}' = \boldsymbol{d} + \boldsymbol{e}_v, \\ u_{\boldsymbol{d}}^{-v} & \text{if } \boldsymbol{d}' = \boldsymbol{d} - \boldsymbol{e}_v, \\ -\mu & \text{if } \boldsymbol{d}' = \boldsymbol{d}, \\ 0 & \text{otherwise,} \end{cases} \tag{4.33}$$

where the cumulative mutation rates are given in equation (4.29).

4.2.2 The four-state model with Hopfield-type fitness

As in the two-state model, in the four-state model the general set-up from the permutation-invariant model shall be adopted with alphabet $\mathcal{A} = \{0, 1, 2, 3\}$ and mutation matrix (3.2). The lumping procedure is very similar to the one used for the two-state model with Hopfield-type fitness. Here again, the pattern $\boldsymbol{\xi}^0$ shall be chosen as $\boldsymbol{\xi}^0 = 00\ldots0$, and the matrix $\boldsymbol{\xi}$ collecting the patterns is of dimension $(p+1) \times N$. However, in the four-state model the matrix $\boldsymbol{\rho}$ collecting the possible column vectors $\boldsymbol{\rho}_v$ has dimension $(p+1) \times 4^p$, and thus the sites are classified into 4^p classes Λ_v. For the case with two patterns, $p = 1$,

the matrix $\boldsymbol{\rho}$ reads explicitly

$$\boldsymbol{\rho} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 3 \end{pmatrix} . \tag{4.34}$$

As in the four-state model the Hamming distance has three components, there are $3 \cdot 4^p$ partial distances $d_{v,k}$ with respect to pattern $\boldsymbol{\xi}^0$, where the first index, v, refers to the subset Λ_v, whereas the second index, k, indicates the type of mutation. The partial distances $d_{v,k}$ are given by

$$d_{v,k} := \sum_{\alpha \in \Lambda_v} \delta(k, \sigma_\alpha) , \tag{4.35}$$

$k \in \{1,2,3\}$, with $d_{v,0} = \sum_{\alpha \in \Lambda_v} \delta(0, \sigma_\alpha) = N_v - \sum_{k=1}^{3} d_{v,k}$ as the number of wild-type sites in Λ_v. The partial distances $d_{v,k}^q$ with respect to pattern $\boldsymbol{\xi}^q$ are given by

$$d_{v,k}^q = \begin{cases} d_{v,k} & \text{if } \rho_v^q = 0, \\ d_{v,0} & \text{if } \rho_v^q = k, \\ d_{v,\ell} & \text{if } \rho_v^q = m \text{ with } \{k,\ell,m\} = \{1,2,3\}. \end{cases} \tag{4.36}$$

For instance, consider the partial distances $d_{v,2}^1$ of type $k = 2$ to pattern $\boldsymbol{\xi}^1$ in the case of two patterns. In subset Λ_1 (with $\boldsymbol{\rho}_1 = (0,0)^T$), the partial distance of type $k = 2$ to pattern $\boldsymbol{\xi}^1$, $d_{1,2}^1$ is given by the partial distance of type $k = 2$ to pattern $\boldsymbol{\xi}^0$, $d_{1,2}^1 = d_{1,2}$. In subset Λ_3 (where $\boldsymbol{\rho}_3 = (0,2)^T$), the partial distance of type $k = 2$ to pattern $\boldsymbol{\xi}^1$ is given by the number of wild-type sites in that partition, and hence $d_{3,2}^1 = d_{3,0}$. In subset Λ_2, $\boldsymbol{\xi}^1$ has entries 1 due to $\boldsymbol{\rho}_2 = (0,1)^T$, and thus the type-2 mutations are those sites that have entries 3, hence $d_{2,2}^1 = d_{2,3}$. Similarly, $d_{4,2}^1 = d_{4,1}$.

The specific distances $\boldsymbol{w}^q = (w_k^q)_{k=1,2,3}$ read

$$w_k^q = \sum_{v=1}^{4^p} d_{v,k}^q . \tag{4.37}$$

Because here again, the specific distances w_k^q can be expressed by the partial distances $d_{v,k}$ with respect to pattern $\boldsymbol{\xi}^0$, these determine the fitness, and therefore the lumping is done with respect to the partial distances $d_{v,k}$, collected in the distance vector $\boldsymbol{d} = (d_{v,k})_{\substack{v=1,\ldots,4^p \\ k=1,2,3}}$.

To get the cumulative mutation rates, each subset Λ_v is considered separately and treated like a four-state model with permutation-invariant fitness. To this end, the definition of the mutational directions $\pm\chi$ for the permutation-invariant four-state model from section 4.1.2 given by k and $+k, -\ell$ with $k, \ell \in \{1,2,3\}$ and $k < \ell$, has to be generalised to $\chi = (v,k)$ or $\chi = (v,+k,-\ell)$ to specify the subset v in question. Hence a mutation in Λ_v can be described by the cumulative mutation rates $u_{\boldsymbol{d}}^{\pm\chi}$, where the sign acts only on the type of mutation specified by k, ℓ, and not on the subset v in question.

A mutation at a site in Λ_v from 0 to k happens at a rate $\mu_k d_{v,0}/N$ and leads to a change $\boldsymbol{d} \to \boldsymbol{d} + \boldsymbol{e}_{v,k}$ with the unit vector of mutation $\boldsymbol{e}_{v,k} = (\delta_{v,w}\delta_{k,\ell})_{\substack{w=1,\ldots,4^p \\ \ell=1,2,3}}$, and similarly for the other types of mutation. Hence the cumulative mutation rates are given by

$$\begin{aligned} u^{v,+k} &= \mu_k d_{v,0}/N && \text{for } \boldsymbol{d} \to \boldsymbol{d} + \boldsymbol{e}_{v,k}, \\ u^{v,-k} &= \mu_k d_{v,k}/N && \text{for } \boldsymbol{d} \to \boldsymbol{d} - \boldsymbol{e}_{v,k}, \\ u^{v,+k,-\ell} &= \mu_m d_{v,\ell}/N && \text{for } \boldsymbol{d} \to \boldsymbol{d} + \boldsymbol{e}_{v,k} - \boldsymbol{e}_{v,\ell}. \end{aligned} \tag{4.38}$$

The number of different $\boldsymbol{d}$ in this model is given by

$$|\mathcal{S}| = \prod_{v=1}^{4^p} [(N_v+1)(N_v+2)(N_v+3)/6] \,, \tag{4.39}$$

and there are

$$n_{\boldsymbol{d}} = \prod_{v=1}^{4^p} \binom{N_v}{d_{v,0},\, d_{v,1},\, d_{v,2},\, d_{v,3}} \tag{4.40}$$

different sequences $\boldsymbol{\sigma}$ mapped onto each $\boldsymbol{d}$. This immediately determines the transformation $\boldsymbol{L}$ that yields the lumped time-evolution operator $\boldsymbol{H}$ like in the previous sections according to equations (4.4), (4.5) and (4.6) with the $n_{\boldsymbol{d}}$ given in equation (4.40), and the lumped mutation matrix is given by

$$M_{\boldsymbol{d}'\boldsymbol{d}} = \begin{cases} u_{\boldsymbol{d}}^{\pm\chi} & \text{if } \boldsymbol{d}' = \boldsymbol{d} \pm \boldsymbol{e}_\chi, \\ -\sum_{k=1}^{3} \mu_k & \text{if } \boldsymbol{d}' = \boldsymbol{d}, \\ 0 & \text{otherwise} \end{cases} \tag{4.41}$$

with the $u_{\boldsymbol{d}}^{\pm\chi}$ from equation (4.38).

4.3 Summary: Common notation in the different models

In this chapter, the original model for sequence evolution has been lumped into an effective model that disregards the particular order of the sequences and therefore works with considerably less types than the original model, making it easier to handle. Depending on how complicated a fitness function is chosen, this lumping can be done with more or less efficiency.

In the case of a two-state model with permutation-invariant fitness function, which completely disregards the order of the sequences, the sequences are lumped into classes that are labelled by the mutational distance d, which is a scalar. For a permutation-invariant four-state model, a three-dimensional mutational distance $\boldsymbol{d} = (d_k)_{k=1,2,3}$ is required.

In a model with Hopfield-type fitness, the distance of sequences to a number $p+1$ of patterns is considered. Here, the N sites are divided into 2^p or 4^p subsets, in the two-state or four-state case, respectively, and in each of these subsets, the subsequences are described by mutational distances like in the permutation-invariant case, such that the mutational distance in the two-state model with Hopfield-type fitness is given by a 2^p-dimensional $\boldsymbol{d} = (d_v)_{v=1,\ldots,2^p}$, and in the four-state model by a $3 \cdot 4^p$-dimensional $\boldsymbol{d} = (d_{v,k})_{\substack{v=1,\ldots,4^p \\ k=1,2,3}}$. The connection between a model with Hopfield-type fitness and one with permutation-invariant fitness is twofold. Firstly, the permutation-invariant fitness can in fact be considered as a special case of the Hopfield-type fitness with only one pattern, $p=0$. Secondly, the subsequences $\boldsymbol{\sigma}_v$ corresponding to the partition induced by the patterns $\boldsymbol{\xi}$ behave effectively permutation-invariant in the sense that the order *within* each subsequence does not influence the fitness.

In the original model for the two-state case, the sequence space $\mathfrak{S}$ contains $|\mathfrak{S}| = 2^N$ types. For a permutation-invariant fitness, these 2^p different sequences are lumped into $|\mathcal{S}| = N+1$ different classes labelled by d, thus the number of types grows linearly instead of exponentially with N. For the four-state permutation-invariant model, the number of

types is reduced from $|\mathfrak{S}| = 4^N$ to $|\mathcal{S}| = (N+1)(N+2)(N+3)/6$, and thus grows only with N^3.

For the Hopfield-type fitness, the number of classes in the lumped system is given by $|\mathcal{S}| = \prod_{v=1}^{2^p}(N_v+1)$ and $|\mathcal{S}| = \prod_{v=1}^{4^p}[(N_v+1)(N_v+2)(N_v+3)/6]$, for the two- and four-state model, respectively. Now consider the case where the patterns are chosen randomly with equal probability for each letter at each site. This results in a multinomial probability distribution for the number of sites N_v in each subset [Whi76],

$$\mathcal{P}(\{N_1,\ldots,N_n\}) = \frac{N!}{n^N \prod_{v=1}^{n} N_v!}, \tag{4.42}$$

where $n = |\mathcal{A}^p|$, thus $n = 2^p$ or $n = 4^p$ in the two- or four-state model, respectively. The means are given by $\bar{N}_v = N/n$ and the variance $\sigma_v^2 = N(n-1)/n^2$ such that $N_v = N/n + \mathcal{O}(\sqrt{N})$. This leads to a scaling with N^{2^p} or $N^{3\cdot 4^p}$, respectively, instead of the behaviour exponential in N as in the original model based on the sequence space $\mathfrak{S}$. Note however, that with a Hopfield-type fitness function, the number of classes scales exponentially with the complexity of the fitness function given by the number of predefined patterns $p+1$.

The original mutation matrix $\mathcal{M}$ containing the single site mutation rates μ/N or μ_k/N, $k = 1,2,3$, is symmetric as forward and back mutation rates are identical in the microscopic model. To describe mutation as a Markov process, the diagonal entries of $\mathcal{M}$ are given by $-\mu$ or $-\sum_{k=1}^{3}\mu_k$, in the two-state or four-state case, respectively. The mutation matrix $\boldsymbol{M}$ governing the lumped process contains the cumulative mutation rates as given in equations (4.3), (4.11), (4.29) and (4.38) as $u_d^{\pm\chi}$, where χ determines the direction of mutation, and the sign specifies forward and backward direction. In the lumped process described by $\boldsymbol{M}$, forward and backward mutation rates differ in general due to different numbers of sequences being lumped together into classes.

In the two-state permutation-invariant case, there is only one direction, $\chi = 1$, whereas in the four-state permutation-invariant case, the possible directions are $\chi = k$ or $\chi = (+k,-\ell)$ with $k,\ell \in \{1,2,3\}$ and $k < \ell$. The latter is required to distinguish between forward and backward mutations. In the Hopfield-type model, for each direction from the

permutation-invariant case, additionally, the subset Λ_v of sites has to be specified, yielding as possible directions in the two-state model $\chi = v$, and in the four-state model $\chi = (v, k)$ and $\chi = (v, +k, -\ell)$. By using the same diagonal entries as for $\mathcal{M}$, the Markov property of $\boldsymbol{M}$ is guaranteed.

CHAPTER 5

The maximum principle

Although the lumping procedure reduces the number of types very efficiently, especially in the simpler permutation-invariant case, the evaluation of the eigenvalues and eigenvectors of the time-evolution operator $\boldsymbol{H}$ still remains a difficult problem for many applications, due to the size of the eigenvalue problem. If one is interested solely in the equilibrium behaviour of the system, however, it is possible to determine the population mean fitness (at least asymptotically for large sequence length N), given by the leading eigenvalue of $\boldsymbol{H}$. This can be done by a simple maximum principle that can be derived from Rayleigh's general maximum principle, which specifies that the leading eigenvalue $\lambda_{\max}$ of a square matrix $\boldsymbol{H}$ of dimension n can be obtained via a maximisation over $\mathbb{R}^n$,

$$\lambda_{\max} = \sup_{\boldsymbol{v}\in\mathbb{R}^n} \frac{\boldsymbol{v}^T\boldsymbol{H}\boldsymbol{v}}{\boldsymbol{v}^T\boldsymbol{v}}. \tag{5.1}$$

The vector $\boldsymbol{v}$ for which the supremum is attained is the eigenvector corresponding to the eigenvalue $\lambda_{\max}$. The simple maximum principle derived from this guarantees that the population mean fitness $\bar{r}$ can be obtained by maximising a function on the mutational distance space $\mathcal{S}$. It can be shown that the maximiser itself is the ancestral mean mutational distance $\hat{x}$.

The results obtained via the maximum principle are exact in three special cases, namely (i) the case of $N \to \infty$, (ii) the case of linear fitness functions and mutation rates, and (iii) the case of unidirectional mutation rates. For any other case, they are often a good approximation. The most important of these special cases is the case $N \to \infty$, because even for rather short sequence lengths, it is usually a good approximation and thus is applicable for any system.

This maximum principle has first been derived for a two-state model with permutation-invariant fitness in [HRWB02], treating all three special cases mentioned above. In this simple version, the mutational distance space $\mathcal{S}$ is a subset of $\mathbb{Z}$, and the mutational distances are given by scalars. The first generalisation of this scalar maximum principle to a case with multidimensional mutational distances was given in [GG04b], treating the four-state model with permutation-invariant fitness for all three special cases. Subsequently, for the case $N \to \infty$, the restriction to a permutation-invariant fitness has been relaxed in [BBBK05], yielding the maximum principle in a general form that is applicable to all models considered in this thesis.

The aim of this chapter is to derive explicitly maximum principles for two- and four-state models with both permutation-invariant and Hopfield-type fitness functions. To this end, in section 5.1, the mutation matrix is symmetrised to provide a common starting point for the case of infinite sequence length and the case of linear fitness functions and mutation rates. In section 5.2, the case of infinite sequence length is considered. For this case, the maximum principle for all models and fitness functions considered here is given by the Theorems 1 and 2 from [BBBK05]. The proof of these theorems shall not be reiterated here. For the original proofs of the two- or four-state model with permutation-invariant fitness, the reader is referred to [HRWB02] or [GG04b], respectively. Section 5.3 contains the proof of the maximum principle for the linear model, whereas the case of unidirectional mutation rates is considered in section 5.4. Although the proofs shown here are not long, they are rather technical.

5.1 Symmetrisation of M

Although the original mutation matrices $\mathcal{M}$ as given in the equations (3.1) and (3.2) for the two- and four-state models are symmetric, with forward and backward mutation rates being identical, the lumped mutation matrices M of equations (4.7), (4.12), (4.33) and (4.41) are non-symmetric. They are however reversible, i.e.,

$$M_{dd'}\pi_{d'} = M_{d'd}\pi_d\,, \tag{5.2}$$

where $\boldsymbol{\pi} = (\pi_{\boldsymbol{d}})_{\boldsymbol{d}\in\mathcal{S}}$ is the stationary distribution of the pure mutation process. This is the equidistribution of types on $\mathfrak{S}$, and thus on $\mathcal{S}$ given by the number of sequences $n_{\boldsymbol{d}}$ that are lumped onto the same mutational distance $\boldsymbol{d}$, $\pi_{\boldsymbol{d}} \sim n_{\boldsymbol{d}}$. The reversibility implies that the mutation matrix $\boldsymbol{M}$ can be symmetrised by the means of a diagonal transformation $\boldsymbol{\Pi} := \mathrm{diag}\{\pi_{\boldsymbol{d}}\}$, which yields the symmetrised mutation matrix as

$$\widetilde{\boldsymbol{M}} = \boldsymbol{\Pi}^{-1/2}\boldsymbol{M}\boldsymbol{\Pi}^{1/2}, \tag{5.3}$$

with off-diagonal entries

$$\widetilde{M}_{\boldsymbol{d}\boldsymbol{d}'} = M_{\boldsymbol{d}\boldsymbol{d}'}\sqrt{\pi_{\boldsymbol{d}'}/\pi_{\boldsymbol{d}}} = M_{\boldsymbol{d}'\boldsymbol{d}}\sqrt{\pi_{\boldsymbol{d}}/\pi_{\boldsymbol{d}'}} = \sqrt{M_{\boldsymbol{d}\boldsymbol{d}'}M_{\boldsymbol{d}'\boldsymbol{d}}} = \widetilde{M}_{\boldsymbol{d}'\boldsymbol{d}}. \tag{5.4}$$

Using the cumulative mutation rates $u_{\boldsymbol{d}}^{\pm\chi}$, this reads

$$\widetilde{M}_{\boldsymbol{d}'\boldsymbol{d}} := \widetilde{u}_{\boldsymbol{d}}^{\pm\chi} = \widetilde{u}_{\boldsymbol{d}\pm\boldsymbol{e}_\chi}^{\mp\chi} = \sqrt{u_{\boldsymbol{d}}^{\pm\chi}u_{\boldsymbol{d}\pm\boldsymbol{e}_\chi}^{\mp\chi}} \quad \text{if } \boldsymbol{d}' = \boldsymbol{d}\pm\boldsymbol{e}_\chi \text{ and 0 otherwise,} \tag{5.5}$$

as can be seen by using the explicit version of the number of sequences $n_{\boldsymbol{d}}$ mapped onto the same mutational distance $\boldsymbol{d}$, given in the equations (4.1), (4.10), (4.32) and (4.40) for permutation-invariant and Hopfield-type fitness in the two- and four-state models, respectively. Because $\boldsymbol{\Pi}$ is diagonal, the diagonal entries of the mutation matrix are unchanged,

$$\widetilde{M}_{\boldsymbol{d}\boldsymbol{d}} = M_{\boldsymbol{d}\boldsymbol{d}} = \begin{cases} -\mu & \text{for the two-state model,} \\ -\sum_{k=1}^{3}\mu_k & \text{for the four-state model.} \end{cases} \tag{5.6}$$

As $\boldsymbol{R}$ is diagonal as well, it is not changed by the transformation $\boldsymbol{\Pi}^{1/2}$, and thus this transformation also symmetrises the time-evolution operator such that

$$\widetilde{\boldsymbol{H}} = \boldsymbol{\Pi}^{-1/2}\boldsymbol{H}\boldsymbol{\Pi}^{1/2} = \boldsymbol{R} + \widetilde{\boldsymbol{M}} \tag{5.7}$$

is symmetric.

Before symmetrisation, $\boldsymbol{H}$ was expressed as the sum of a Markov generator $\boldsymbol{M}$ and a diagonal remainder $\boldsymbol{R}$. As the transformation $\boldsymbol{\Pi}$ does not preserve the Markov property, this is not the case for the symmetrised time-evolution operator in (5.7). It is however useful to split it up this way. To this end, let

$$\widetilde{\boldsymbol{H}} = \boldsymbol{E} + \boldsymbol{F}, \tag{5.8}$$

where $\boldsymbol{F}$ is a (symmetric) Markov generator and $\boldsymbol{E}$ is the (diagonal) remainder. The off-diagonal entries of $\boldsymbol{F}$ are given by those of $\widetilde{M}$ (5.5), $F_{\boldsymbol{d}'\boldsymbol{d}} = \widetilde{M}_{\boldsymbol{d}'\boldsymbol{d}}$ for $\boldsymbol{d}' \neq \boldsymbol{d}$, whereas the Markov property requires as diagonal entries

$$F_{\boldsymbol{d}\boldsymbol{d}} = -\sum_{\boldsymbol{d}'} F_{\boldsymbol{d}'\boldsymbol{d}} = -\sum_{\chi} \left(\widetilde{u}_{\boldsymbol{d}}^{+\chi} + \widetilde{u}_{\boldsymbol{d}}^{-\chi} \right) = -\sum_{\chi} \left(\sqrt{u_{\boldsymbol{d}}^{+\chi} u_{\boldsymbol{d}+\boldsymbol{e}_\chi}^{-\chi}} + \sqrt{u_{\boldsymbol{d}}^{-\chi} u_{\boldsymbol{d}-\boldsymbol{e}_\chi}^{+\chi}} \right) \tag{5.9}$$

The remainder $\boldsymbol{E}$ is given by

$$\begin{aligned} E_{\boldsymbol{d}} &= R_{\boldsymbol{d}} + \widetilde{M}_{\boldsymbol{d}\boldsymbol{d}} - F_{\boldsymbol{d}\boldsymbol{d}} \\ &= R_d - \sum_{\chi} \left(u_{\boldsymbol{d}}^{+\chi} + u_{\boldsymbol{d}}^{-\chi} \right) + \sum_{\chi} \left(\sqrt{u_{\boldsymbol{d}}^{+\chi} u_{\boldsymbol{d}+\boldsymbol{e}_\chi}^{-\chi}} + \sqrt{u_{\boldsymbol{d}}^{-\chi} u_{\boldsymbol{d}-\boldsymbol{e}_\chi}^{+\chi}} \right) \end{aligned} \tag{5.10}$$

5.2 The limit of infinite sequence length

In the case of infinite sequence length, the maximum principle has been shown to hold exactly for a two-state model with permutation-invariant fitness; this has been generalised to a four-state model with permutation-invariant fitness in [GG04b] and finally generalised even further to include also a Hopfield-type fitness in [BBBK05]. For finite sequence lengths N, it is exact up to $\mathcal{O}\left(\frac{1}{N}\right)$.

To deal with the case of infinite sequence length, it will prove useful to use intensively scaled normalised versions of the extensively scaled variables like the mutational distances. The pattern in the Hopfield model, previously characterised by the *number* of sites N_v in each subset Λ_v, will now be described by the *fraction* of sites in Λ_v, given by $X_v := N_v/N$.

Similarly, we use normalised mutational distances

$$x_v := d_v/N_v\,, \tag{5.11}$$

with $x_v \in [0,1]$ for the two-state model, and for the four-state model

$$x_{v,k} := d_{v,k}/N_v \tag{5.12}$$

with $0 \leq x_{v,k}$ and $\sum_{k=1}^{3} x_{v,k} \leq 1$. Furthermore, the fraction of wild-type sites in subset Λ_v is given by $x_{v,0} = 1 - \sum_{k=1}^{3} x_{v,k}$. The permutation-invariant model is obtained in the case $p = 0$. For the x_v or $x_{v,k}$, the scaling is chosen differently for each Λ_v (namely with

N_v rather than with N), because this yields values of $x_v \in [0,1]$ or $\sum_{k=1}^{3} x_{v,k} \in [0,1]$, respectively, independently of the pattern determined by the N_v. Furthermore, this is the convention used in the literature on the Hopfield model.

For finite N, $\boldsymbol{x}$ takes rational values in a normalised version of the mutational distance space $\frac{1}{N} \cdot \mathcal{S} \subset \mathcal{D}$, where $\mathcal{D}$ is a compact domain in $\mathbb{R}^n$, with $n = 2^p$ or $n = 3 \cdot 4^p$ for the two- or four-state model, respectively. For $N \to \infty$, the vectors $\boldsymbol{x}$ become dense in $\mathcal{D}$.

Assume that the entries of $\widetilde{\boldsymbol{H}} = \boldsymbol{E} + \boldsymbol{F}$ can be approximated by functions e and f from $C_b^2(\mathcal{D}, \mathbb{R})$, i.e., twice continuously differentiable functions with bounded second derivatives that map $\mathcal{D}$ onto $\mathbb{R}$ such that

$$E_{\boldsymbol{d}} = e(\boldsymbol{x}_{\boldsymbol{d}}) + \mathcal{O}\left(\frac{1}{N}\right), \tag{5.13}$$

$$F_{\boldsymbol{d}'\boldsymbol{d}} = f_{\boldsymbol{\Delta}}(\boldsymbol{x}_{\boldsymbol{d}}) + \mathcal{O}\left(\frac{1}{N}\right), \tag{5.14}$$

where $\boldsymbol{\Delta} = \boldsymbol{d}' - \boldsymbol{d}$ and the notation $\boldsymbol{x}_{\boldsymbol{d}}$ is used to emphasise that the normalised mutational distance $\boldsymbol{x}$ corresponding to a particular $\boldsymbol{d}$ is meant. Then Theorem 1 from [BBBK05] applies here, giving the leading eigenvalue $\lambda_{\max}$ of $\widetilde{\boldsymbol{H}}$, which coincides with the population mean fitness in equilibrium, $\bar{r} = \lambda_{\max}$. For $\mathcal{S} \subset \mathbb{Z}^n$, the slightly abusive notation $\mathcal{S} - \boldsymbol{d} := \{\boldsymbol{d}' - \boldsymbol{d} | \boldsymbol{d}' \in \mathcal{S}\}$ shall be used.

Theorem 1 (from [BBBK05]). *Assume that $E_{\boldsymbol{d}}$ and $F_{\boldsymbol{d}'\boldsymbol{d}}$ are as in equations (5.13) and (5.14). Assume further that the $C_b^2(\mathcal{D}, \mathbb{R})$ function e assumes its absolute maximum in the interior* $\mathrm{int}(\mathcal{D})$*, and that f satisfies*

$$\sum_{\boldsymbol{\Delta} \in \mathcal{S} - \boldsymbol{d}} f_{\boldsymbol{\Delta}}(\boldsymbol{x}_{\boldsymbol{d}}) |\Delta_w| \Delta_{w'}^2 \leq C \tag{5.15}$$

for some constant C, uniformly for all $\boldsymbol{d} \in \mathcal{S}$ and $1 \leq w, w' \leq n$. Then, there exist constants $0 \leq C', C'' < \infty$ such that

$$e(\boldsymbol{x}^*) - \frac{C'}{N} \leq \lambda_{\max} \leq e(\boldsymbol{x}^*) + \frac{C''}{N}, \tag{5.16}$$

where $\boldsymbol{x}^*$ is a point where $e(\boldsymbol{x})$ assumes its maximum.

Remark. For the proof of Theorem 1, it is actually sufficient that the functions e and

f_Δ are C_b^2 locally, in a neighbourhood of $\boldsymbol{x}^*$. This means that also fitness functions with finitely many jumps are possible.

As the mutation rates are given explicitly, for $\boldsymbol{F}$, the validity of assumption (5.14) can immediately be verified. Let the functions $f_\Delta(\boldsymbol{x})$ be given by

$$f_\Delta(\boldsymbol{x}) = \begin{cases} \widetilde{u}^\chi(\boldsymbol{x_d}) & \text{if } \boldsymbol{\Delta} = \boldsymbol{e}_\chi, \\ -2\sum \widetilde{u}^\chi(\boldsymbol{x_d}) & \text{if } \boldsymbol{\Delta} = 0, \\ 0 & \text{otherwise.} \end{cases} \tag{5.17}$$

The functions $\widetilde{u}^\chi(\boldsymbol{x})$ are given by

$$\widetilde{u}^\chi(\boldsymbol{x}) := \sqrt{u^{+\chi}(\boldsymbol{x})u^{-\chi}(\boldsymbol{x})}\,, \tag{5.18}$$

with the cumulative mutation rates $u^{\pm\chi}(\boldsymbol{x_d}) = u_{\boldsymbol{d}}^{\pm\chi}$ from equations (4.3), (4.11), (4.29) and (4.38), respectively. Thus they read explicitly for the two-state model

$$u^+(x) = \mu(1-x) \qquad u^-(x) = \mu x \tag{5.19}$$

$$u^{+v}(\boldsymbol{x}) = \mu X_v(1-x_v) \qquad u^{-v}(\boldsymbol{x}) = \mu X_v x_v \tag{5.20}$$

with permutation-invariant fitness (5.19) and Hopfield-type fitness (5.20), and for the four-state model

$$u^{+k}(\boldsymbol{x}) = \mu_k x_0 \qquad u^{-k}(\boldsymbol{x}) = \mu_k x_k \qquad u^{+k,-\ell}(\boldsymbol{x}) = \mu_m x_\ell \tag{5.21}$$

$$u^{v,+k}(\boldsymbol{x}) = \mu_k X_v x_{v,0} \qquad u^{v,-k}(\boldsymbol{x}) = \mu_k X_v x_{v,k} \qquad u^{v,+k,-\ell}(\boldsymbol{x}) = \mu_m X_v x_{v,\ell} \tag{5.22}$$

with $\{k,\ell,m\} = \{1,2,3\}$, for permutation-invariant fitness (5.21) and Hopfield-type fitness (5.22). Using a Taylor approximation, it can be shown that the differences between the exact entries of F as given in equations (5.5) and (5.9), and their approximations from equation (5.17), are indeed of $\mathcal{O}\left(\frac{1}{N}\right)$.

Equation (5.15) is a condition about how fast the off-diagonal entries of $\widetilde{\boldsymbol{H}}$ have to decay, when one moves further away from the diagonal. The f_Δ from equation (5.17) fulfil this condition, because they are bounded functions, and for each $\boldsymbol{d}$ there are only finitely

many $\boldsymbol{\Delta}$ with $f_{\boldsymbol{\Delta}} \neq 0$, given by the unit vectors of mutation $\pm \boldsymbol{e}_\chi$, and finally the possible entries Δ_w of $\boldsymbol{\Delta}$ are limited to $0, \pm 1$.

Assuming that also the reproduction rates $R_{\boldsymbol{d}}$ can be approximated by a C_b^2 function $r(\boldsymbol{x})$ as

$$R_{\boldsymbol{d}} = r(\boldsymbol{x}_{\boldsymbol{d}}) + \mathcal{O}\left(\frac{1}{N}\right), \tag{5.23}$$

then from equation (5.10), the matrix $\boldsymbol{E}$ is approximated by

$$e(\boldsymbol{x}) = r(\boldsymbol{x}) - \sum_{\chi} \left(u^{+\chi}(\boldsymbol{x}) + u^{-\chi}(\boldsymbol{x}) - 2\sqrt{u^{+\chi}(\boldsymbol{x}) u^{-\chi}(\boldsymbol{x})} \right), \tag{5.24}$$

fulfilling equation (5.13). With the definition of the mutational loss function as

$$g(\boldsymbol{x}) := \sum_{\chi} \left(u^{+\chi}(\boldsymbol{x}) + u^{-\chi}(\boldsymbol{x}) - 2\sqrt{u^{+\chi}(\boldsymbol{x}) u^{-\chi}(\boldsymbol{x})} \right) \tag{5.25}$$

i.e.,

$$g(\boldsymbol{x}) = \mu \sum_{v=1}^{2^p} X_v \left[1 - 2\sqrt{x_v (1 - x_v)} \right] \tag{5.26}$$

and

$$g(\boldsymbol{x}) = \sum_{v=1}^{4^p} \sum_{k=1}^{3} \mu_k X_v \left[1 - 2\sqrt{x_{v,0} x_{v,k}} - 2\sqrt{x_{v,\ell} x_{v,m}} \right] \tag{5.27}$$

with $\{k, \ell, m\} = \{1, 2, 3\}$, for the two- and four-state model, respectively, this reads explicitly

$$e(\boldsymbol{x}) = r(\boldsymbol{x}) - g(\boldsymbol{x}), \tag{5.28}$$

and thus equation (5.16) can be written for the case considered here as

$$\bar{r} = \sup_{\boldsymbol{x} \in \mathcal{D}} \left(r(\boldsymbol{x}) - g(\boldsymbol{x}) \right) + \mathcal{O}\left(\frac{1}{N}\right). \tag{5.29}$$

Theorem 2 of [BBBK05] is concerned with the localisation of the ancestral distribution $\boldsymbol{a}$ (from section 2.1).

Theorem 2 (from [BBBK05]). *Let $E_{\boldsymbol{d}}$ and $F_{\boldsymbol{d}'\boldsymbol{d}}$ satisfy the hypotheses of Theorem 1. Assume that e assumes its maximum at a unique point $\boldsymbol{x}^* \in \mathrm{int}(\mathcal{D})$, and that the Hessian of e at $\boldsymbol{x}^*$ is negative definite. Then, for every $0 < \beta \leq 1$, there is a $\rho > 0$, independent*

of N, so that, for N large enough:

$$\sum_{\substack{d\in\mathcal{S}:\\ |\varpi_d-\varpi^*|\geq\sqrt{\rho/\beta N}}} a_d \leq \beta\,, \tag{5.30}$$

where $\boldsymbol{a} = (a_d)$ is the ancestral distribution.

This means that the ancestral distribution is concentrated around the maximum $\boldsymbol{x}^*$ with a width of the order of $1/\sqrt{N}$. In the limit of $N \to \infty$, the ancestral distribution approaches a point measure located at $\boldsymbol{x}^*$. Hence, in this limit the ancestral mean mutational distance converges toward the maximum position, $\hat{\boldsymbol{x}} = \sum_{d\in\mathcal{S}} a_d \boldsymbol{x}_d \to \boldsymbol{x}^*$.

Let us now summarise the results of Theorems 1 and 2 applied to the mutation–selection model used here.

Theorem 3 (The maximum principle for $N \to \infty$). *(i) Assume that for the lumped mutation–selection model as set up in chapter 4 it is possible to approximate the reproduction rates R_d by a C_b^2 function $r(\boldsymbol{x}_d)$ as specified in equation (5.23), and that the C_b^2 function e assumes its global maximum in* $\mathrm{int}(\mathcal{D})$. *Then the population mean fitness in equilibrium is given by*

$$\bar{r} = \sup_{\boldsymbol{x}\in\mathcal{D}} [r(\boldsymbol{x}) - g(\boldsymbol{x})] + \mathcal{O}\left(\frac{1}{N}\right). \tag{5.31}$$

(ii) Assume furthermore that e assumes its maximum at a unique point $\boldsymbol{x}^*$ in $\mathrm{int}(\mathcal{D})$, *and that the Hessian of e at $\boldsymbol{x}^*$ is negative definite. Then in the limit of $N \to \infty$, the maximiser $\boldsymbol{x}^*$ is given by the mean ancestral mutational distance $\hat{\boldsymbol{x}}$, and in particular*

$$\bar{r} = r(\hat{\boldsymbol{x}}) - g(\hat{\boldsymbol{x}})\,. \tag{5.32}$$

Proof. This follows from Theorems 1 and 2, which have been shown to apply in the two- and four-state models with both permutation-invariant and Hopfield-type fitness functions in the preceding discussion. In particular, the cumulative mutation rates as given in equations (5.19) to (5.22) have been shown to fulfil condition (5.15) from Theorem 1.

Remark. Note that for the model with Hopfield-type fitness in the limit of $N \to \infty$, it is not required that all $N_v \to \infty$. Should an N_v remain finite in the limit, then $X_v \to 0$ and thus the corresponding subset of sites Λ_v does not contribute to the properties of the system.

5.3 The linear model

In the linear model, it is assumed that both the fitness function and the mutation rates depend linearly on functions $y_{v,k}$ of the genotype components $x_{v,k}$, i.e., they are sums of functions that depend on only one of the $x_{v,k}$ each, and thus can be written as

$$r(\boldsymbol{x}) = r_0 - \sum_{v,k} \alpha_{v,k} y_{v,k}(x_{v,k}) \quad \text{and} \quad u^{\pm\chi}(\boldsymbol{x}) = u_0^{\pm\chi} + \sum_{v,k} \beta_{v,k}^{\pm\chi} y_{v,k}(x_{v,k}) \tag{5.33}$$

with parameters $\alpha_{v,k}$ and $\beta_{v,k}^{\pm\chi}$, where for the two-state model $v = 1, \ldots 2^p$ and $k = 1$ and thus the summation over k is skipped. In the four-state model, $v = 1, \ldots, 4^p$ and $k = 1, 2, 3$. The permutation-invariant case is again obtained with $p = 0$, i.e., $v = 1$, and thus skipping the sum over v. Note that in equation (5.33) the same functions $y_{v,k}$ occur in both fitness function and mutation rates. The simplest and most common case is the case where $y_{v,k}(x_{v,k}) = x_{v,k}$, i.e., a directly linear model. This type of model has been used e.g., in [vHB86, GH02] in a permutation-invariant two-state version. As the cumulative mutation rates given in equations (5.19) to (5.22) are linear, one is in the linear set-up when choosing a linear fitness function. Then the following applies:

Theorem 4 (The maximum principle for the linear model). *Let the fitness function and mutation rates be given as in equation (5.33). Then the population mean fitness is given by*

$$\bar{r} = \sup_{\boldsymbol{x} \in \mathcal{D}} [r(\boldsymbol{y}(\boldsymbol{x})) - g(\boldsymbol{y}(\boldsymbol{x}))] = r(\hat{\boldsymbol{y}}) - g(\hat{\boldsymbol{y}}) \,. \tag{5.34}$$

Remark. In the case of a directly linear model, $y_{v,k}(x_{v,k}) = x_{v,k}$, which is the one

resulting from mutation model considered here, this simplifies to

$$\bar{r} = \sup_{\boldsymbol{x}\in\mathcal{D}} [r(\boldsymbol{x}) - g(\boldsymbol{x})] = \hat{r} - g(\hat{\boldsymbol{x}})\,. \tag{5.35}$$

For the two-state model with permutation-invariant fitness, this result has been shown in [HRWB02]. The generalisation to a multidimensional mutational distance for a four-state model with permutation-invariant fitness was obtained in [GG04b], and this proof is applicable with minimal changes also for the Hopfield-type fitness.

Proof. The starting point for the derivation of the maximum principle is an eigenvalue equation for the symmetrised time-evolution operator $\widetilde{\boldsymbol{H}}$. From equation (5.7) it is evident that both time-evolution operators, $\boldsymbol{H}$ and $\widetilde{\boldsymbol{H}}$ have the same eigenvalues, but the eigenvectors differ. The left and right eigenvectors $\widetilde{\boldsymbol{z}}$ and $\widetilde{\boldsymbol{p}}$ of $\widetilde{\boldsymbol{H}}$ for the largest eigenvalue $\bar{r}$ are related to the corresponding eigenvectors $\boldsymbol{z}$ and $\boldsymbol{p}$ of $\boldsymbol{H}$ by

$$\widetilde{\boldsymbol{z}} = \boldsymbol{z}\,\boldsymbol{\Pi}^{1/2} \quad \text{and} \quad \widetilde{\boldsymbol{p}} = \boldsymbol{\Pi}^{-1/2}\boldsymbol{p} \tag{5.36}$$

with the diagonal transformation $\boldsymbol{\Pi}$ from equation (5.3). The relation between the ancestral distribution $\boldsymbol{a}$ and the symmetrised population $\widetilde{\boldsymbol{p}}$ is given by

$$a_i = z_i\,p_i = (\widetilde{\boldsymbol{z}}\,\boldsymbol{\Pi}^{-1/2})_i(\boldsymbol{\Pi}^{1/2}\widetilde{\boldsymbol{p}})_i = \widetilde{z}_i\,\widetilde{p}_i = c\cdot\widetilde{p}_i^{\,2}\,, \tag{5.37}$$

with the proportionality constant c being independent of the type i, because $\widetilde{\boldsymbol{z}} \propto \widetilde{\boldsymbol{p}}$ due to the symmetry of $\widetilde{\boldsymbol{H}}$. With the relation $\widetilde{\boldsymbol{p}} \propto \sqrt{\boldsymbol{a}}$, the eigenvalue equation of $\widetilde{\boldsymbol{H}}$ in ancestral formulation becomes $\widetilde{\boldsymbol{H}}\sqrt{\boldsymbol{a}} = \bar{r}\sqrt{\boldsymbol{a}}$, which reads explicitly

$$\begin{aligned}\bar{r}\sqrt{a_{\boldsymbol{d}}} = &\left(r_{\boldsymbol{d}} - \sum_{\chi}\left(u_{\boldsymbol{d}}^{+\chi} + u_{\boldsymbol{d}}^{-\chi}\right)\right)\sqrt{a_{\boldsymbol{d}}} \\ &+ \sum_{\chi}\left(\sqrt{u_{\boldsymbol{d}-\boldsymbol{e}_\chi}^{+\chi}u_{\boldsymbol{d}}^{-\chi}}\sqrt{a_{\boldsymbol{d}-\boldsymbol{e}_\chi}} + \sqrt{u_{\boldsymbol{d}+\boldsymbol{e}_\chi}^{-\chi}u_{\boldsymbol{d}}^{+\chi}}\sqrt{a_{\boldsymbol{d}+\boldsymbol{e}_\chi}}\right)\,,\end{aligned} \tag{5.38}$$

for some value of the distance $\boldsymbol{d}$. For the $a_{\boldsymbol{d}}$ an ansatz is made such that

$$\frac{a_{\boldsymbol{d}+\boldsymbol{e}_\chi}}{a_{\boldsymbol{d}}} = C_\chi\frac{u_{\boldsymbol{d}}^{+\chi}}{u_{\boldsymbol{d}+\boldsymbol{e}_\chi}^{-\chi}}\,. \tag{5.39}$$

This is consistent for all directions of mutation χ due to the reversibility of $\boldsymbol{M}$ (5.2). To determine the constants C_χ, equation (5.39) is multiplied by its denominators and summed

over all $\boldsymbol{d}$, which yields the ancestral means of the mutation rates

$$\sum_{\boldsymbol{d}} u_{\boldsymbol{d}+\boldsymbol{e}_\chi}^{-\chi} a_{\boldsymbol{d}+\boldsymbol{e}_\chi} = C_\chi \sum_{\boldsymbol{d}} u_{\boldsymbol{d}}^{+\chi} a_{\boldsymbol{d}} \quad \Longleftrightarrow \quad C_\chi = \frac{\hat{u}^{-\chi}}{\hat{u}^{+\chi}} . \tag{5.40}$$

Now, equation (5.38) is divided by $\sqrt{a_{\boldsymbol{d}}}$, and the ansatz (5.39) is inserted,

$$\bar{r} = r_{\boldsymbol{d}} - \sum_{\chi} \left(u_{\boldsymbol{d}}^{+\chi} + u_{\boldsymbol{d}}^{-\chi} \right) + \sum_{\chi} \left(\frac{1}{\sqrt{C_\chi}} \, u_{\boldsymbol{d}}^{-\chi} + \sqrt{C_\chi} \, u_{\boldsymbol{d}}^{+\chi} \right) \tag{5.41}$$

Multiplication by $a_{\boldsymbol{d}}$ and summation over all $\boldsymbol{d}$ yields, using the explicit form (5.40) of the C_χ,

$$\bar{r} = \hat{r} - \sum_{\chi} \left(\hat{u}^{+\chi} + \hat{u}^{-\chi} - 2\sqrt{\hat{u}^{+\chi}\hat{u}^{-\chi}} \right) . \tag{5.42}$$

So far, linearity has not been used. With the definition of the mutational loss function from equation (5.25) and $\hat{r} = r(\hat{\boldsymbol{y}})$ due to linearity of r, this yields

$$\bar{r} = r(\hat{\boldsymbol{y}}) - g(\hat{\boldsymbol{y}}) , \tag{5.43}$$

the right part of the maximum principle (5.34).

In order to obtain the supremum condition of (5.34), consider equation (5.41) for two different sequences $\boldsymbol{d}$ and $\boldsymbol{d}'$ and take the difference, using the explicit representation of fitness and mutation functions given in equation (5.33),

$$\begin{aligned} 0 = \sum_{v,k} \Bigg[-\alpha_{v,k} - \sum_{\chi} \left(\beta_{v,k}^{+\chi} + \beta_{v,k}^{-\chi} - \sqrt{C_\chi}\beta_{v,k}^{+\chi} - \frac{\beta_{v,k}^{-\chi}}{\sqrt{C_\chi}} \right) \Bigg] \\ \times \left(y_{v,k}(x_{v,k}) - y_{v,k}(x'_{v,k}) \right) . \end{aligned} \tag{5.44}$$

This is just the condition

$$0 = \sum_{v,k} \left[\frac{\partial}{\partial y_{v,k}} \left[r(\boldsymbol{y}) - g(\boldsymbol{y}) \right]_{\boldsymbol{y}=\hat{\boldsymbol{y}}} \right] \left(y_{v,k}(x_{v,k}) - y_{v,k}(x'_{v,k}) \right) , \tag{5.45}$$

which has to be fulfilled for arbitrary $\boldsymbol{x}$ and $\boldsymbol{x}'$. Hence,

$$0 = \frac{\partial}{\partial y_{v,k}} \left[r(\boldsymbol{y}) - g(\boldsymbol{y}) \right]_{\boldsymbol{y}=\hat{\boldsymbol{y}}} \quad \forall v,k . \tag{5.46}$$

This is a necessary condition for the existence of an extremum at $\hat{\boldsymbol{y}}$. A sufficient condition for the existence of a maximum of the function $r - g$ in $\hat{\boldsymbol{y}}$ is that the Hessian

$$\mathfrak{H}_{(v,k)(w,\ell)}(\hat{\boldsymbol{y}}) := \left[\frac{\partial^2 \left(r(\boldsymbol{y}) - g(\boldsymbol{y}) \right)}{\partial y_{v,k} \partial y_{w,\ell}} \right]_{\boldsymbol{y}=\hat{\boldsymbol{y}}} \tag{5.47}$$

of the second derivatives in the point $\hat{\boldsymbol{y}}$ is a negative definite matrix. The Hessian can be written in the form

$$\mathfrak{H}(\boldsymbol{y}) = -\sum_{\chi} c_{\chi}(\boldsymbol{y}) \boldsymbol{U}_{\chi}(\boldsymbol{y}) \boldsymbol{U}_{\chi}^{T}(\boldsymbol{y}), \tag{5.48}$$

with $c_{\chi}(\boldsymbol{y}) = \frac{1}{2}\left(u^{+\chi}(\boldsymbol{y})\, u^{-\chi}(\boldsymbol{y})\right)^{-3/2}$ and the column vector $\boldsymbol{U}_{\chi}(\boldsymbol{y}) = (U_{\chi,v,k})$ with $U_{\chi,v,k}(\boldsymbol{y}) = \beta_{v,k}^{+\chi} u^{-\chi}(\boldsymbol{y}) - \beta_{v,k}^{-\chi} u^{+\chi}(\boldsymbol{y})$. To test the Hessian for negative definiteness, evaluate the quadratic form for an arbitrary vector $\boldsymbol{z}$,

$$\boldsymbol{z}^{T} \mathfrak{H}\, \boldsymbol{z} = -\sum_{\chi} c_{\chi}(\boldsymbol{y}) \left(\sum_{v,k} z_k U_{\chi,v,k}(\boldsymbol{y}) \right)^{2}, \tag{5.49}$$

which is non-positive for all $\boldsymbol{z}$, and generically negative unless all terms in the sum vanish. Hence, there is a maximum at $\hat{\boldsymbol{y}}$. As the Hessian is negative definite on the whole mutational distance space, there can be no minima or saddle points. Hence, the maximum in $\hat{\boldsymbol{y}}$ is the only maximum, because two different maxima must be divided by saddle points or minima, and therefore it is the global maximum. This proves the extremum condition of the maximum principle (5.34). □

5.4 Unidirectional mutation rates

Up to now, only models where the underlying mutation rates for the process on $\mathfrak{S}$ in forward and backward direction are the same were considered. The extreme case of direction-dependent mutation is the case of unidirectional mutation, meaning that only such mutations that increase the total mutational distance $d = \sum_{v,k} d_{v,k}$ happen, the mutation rates toward types with smaller or equal d are zero, i.e.,

$$u^{+v}(\boldsymbol{x}) > 0 \qquad u^{-v}(\boldsymbol{x}) = 0, \tag{5.50}$$

$$u^{v,+k}(\boldsymbol{x}) > 0 \qquad u^{v,-k}(\boldsymbol{x}) = u^{v,+k,-\ell} = 0, \tag{5.51}$$

for the two- and four-state model, respectively, with distinct $k, \ell \in \{1, 2, 3\}$. This mutation scheme does not go in line with the mutation rates for the DNA system as given in equations (5.19) to (5.22).

In the permutation-invariant model, however, unidirectional mutation rates still are a reasonable approximation. If in the DNA model, selection is sufficiently strong and the fitness function is monotonically decreasing, such that the reference sequence is in fact the sequence with maximal fitness, most individuals present in the population will have a genotype with a small mutational distance from the wild-type. The mutations that leave d constant (they only occur in the four-state model, $\boldsymbol{e}_\chi = \boldsymbol{e}_\ell - \boldsymbol{e}_k$), and those that decrease d ($\boldsymbol{e}_\chi = -\boldsymbol{e}_k$), which in the case of a decreasing fitness function are neutral and advantageous, respectively, occur with rates proportional to x_k, and therefore are rare for individuals with small mutational distances $x = \sum_k x_k = d/N$, which form the main part of the population. In contrast, the mutations that increase d ($\boldsymbol{e}_\chi = \boldsymbol{e}_k$) happen with rates proportional to $1 - x$, which are of order 1 for small x. Therefore, it is reasonable to approximate the mutation rates of the DNA model by unidirectional mutation rates, i.e., $u_{\boldsymbol{d}}^- = 0$ or $u_{\boldsymbol{d}}^{-k} = u_{\boldsymbol{d}}^{\pm k, \mp \ell} = 0$. This is the well known infinite sites limit (see [Kim69], [Ewe79, chapter 8]).

Unidirectional mutation rates have to be defined with respect to one sequence; in the permutation-invariant model this is the wild-type. In the Hopfield-type fitness, there are however a number of patterns, none of which should be considered as special, and thus the notion of unidirectional mutation rates is less meaningful. This is why here only the cases of permutation-invariant fitness shall be treated, although the method for the four-state model generalises immediately to the Hopfield-type fitness.

For unidirectional mutation, the mutation matrix $\boldsymbol{M}$ is no longer irreducible. Hence, in this case, the Perron-Frobenius theorem does not apply. The equilibrium is not unique, but depends on initial conditions. Once the wild-type is lost in the population, it can never occur again, because the mutation rates back to types with smaller d are zero.

The two-state model with permutation-invariant fitness is treated in [HRWB02], and this has been generalised to the four-state model in [GG04b].

5.4.1 The two-state model with permutation-invariant fitness

As the mutation matrix is not irreducible for unidirectional mutation, the population distribution in equilibrium is not unique, but depends on the initial conditions. The equilibrium with the highest mean fitness is always assumed if the wild-type initially occurs with non-zero frequency, or if one considers the case of unidirectional mutation rates as an approximation to the limit of small, but non-vanishing back mutations.

The eigenvalue equation for $\boldsymbol{H}$ in the two-state model with permutation-invariant fitness and unidirectional mutation takes the simple form

$$(\bar{r} - \lambda_d)p_d = u_{d-1}^{+} p_{d-1} \quad \text{with} \quad \lambda_d = r_d - g(x_d) = r_d - u_d^{+}\,, \tag{5.52}$$

where the diagonal entries are the eigenvalues λ_d, because $\boldsymbol{H}$ is a lower triangular matrix. In this case, the situation is particularly simple, and some properties of the population distribution $\boldsymbol{p}$ can be inferred directly, which allow to show that the maximum principle holds in this case in the following form.

Theorem 5 (The maximum principle for the two-state model with unidirectional mutation rates). *Let the mutation rates be as in equation (5.50), and let $d_{\min}$ be the minimal mutational distance that is present in the initial population $\boldsymbol{p}(0)$. Then, the equilibrium population mean fitness is given by*

$$\bar{r} = \sup_{d \geq d_{\min}} [r(x_d) - g(x_d)]\,. \tag{5.53}$$

If, additionally, the maximum in (5.53) is taken at a unique mutational distance $\hat{d}$, it is furthermore

$$\bar{r} = r(\hat{x}) - g(\hat{x})\,, \tag{5.54}$$

where $\hat{x} = x_{\hat{d}}$.

Proof. Consider first the case where the wild-type is initially present in the population, i.e., $d_{\min} = 0$. Suppose there is one d^* such that $p_{d^*} > 0$, but $p_{d^*-1} = 0$. This can only happen if $\bar{r} = \lambda_{d^*}$, and it directly yields that $p_d = 0$ for all $d < d^*$. Indeed, if there was a

sequence $d^+ > d^*$ with $\lambda_{d^+} > \lambda_{d^*}$, p_{d^+} would be negative, in contradiction to the condition that $p_d \geq 0$ for all d. If $\lambda_{d^+} = \lambda_{d^*}$, $p_d = 0$ for all $d < d^+$, including d^*. Thus in this case, the population consists only of the sequences with $d \geq d^+$. This yields the first part of the maximum principle (5.53) with $d_{\min} = 0$.

If, however, the wild-type is not present in the initial population $\boldsymbol{p}(0)$, only the part of the mutational distance space with $d \geq d_{\min}$ has to be considered, where $d_{\min}$ is the minimal mutational distance that is present initially, $p_{d_{\min}}(0) > 0$. Now, the mean fitness assumes the highest possible value in this subspace, which is assumed for at least one sequence d^*, yielding the maximum principle (5.53) for arbitrary initial conditions.

If this maximum is unique, the equilibrium population is given by the right eigenvector corresponding to this eigenvalue λ_{d^*}, which has non-zero entries only for sequences $d \geq d^*$. The left eigenvector corresponding to λ_{d^*} has non-zero entries only for sequences $d \leq d^*$, so that the only d with $a_d \neq 0$ is d^*, and hence $d^* = \hat{d}$ is the only ancestor. This yields the second part of the maximum principle (5.54). □

5.4.2 The four-state model with permutation-invariant fitness

The argument in the four-state model goes very much along the lines of the two-state model, but due to the more complex structure of the mutational distance space (being a subset of $\mathbb{Z}^3$ rather than $\mathbb{Z}$), there are additional points to consider. The hierarchical structure of the mutational distance space in the four-state model is shown in figure 5.1, where the wild-type on the top corner of the mutational distance space "feeds" all mutants underneath.

The mutational distance space $\mathcal{S}$ can be divided into three domains with respect to each sequence $\boldsymbol{d}$, namely the ancestral cone, the offspring cone and the sibling domain. Here, all sequences that can mutate to $\boldsymbol{d}$ lie in the ancestral cone $AC(\boldsymbol{d})$, all sequences to which $\boldsymbol{d}$ can mutate lie in the offspring cone $OC(\boldsymbol{d})$, whereas the sequences that are not connected to $\boldsymbol{d}$ via a mutational path form the sibling domain $SD(\boldsymbol{d})$, which is the remainder of the mutational distance space.

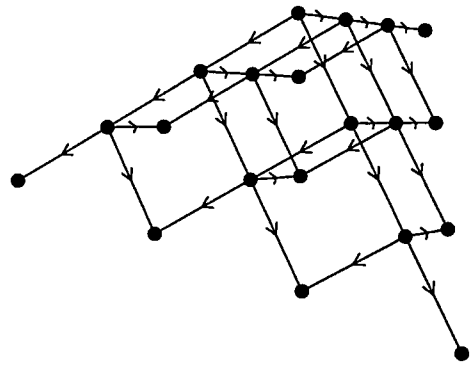

Figure 5.1: Structure of the mutational distance space S for the four-state model with unidirectional mutation.

The eigenvalue equation of $\boldsymbol{H}$ for unidirectional mutation in the four-state model is given by

$$(\bar{r} - \lambda_{\boldsymbol{d}})\, p_{\boldsymbol{d}} = \sum_{k=1}^{3} u_{\boldsymbol{d}-\boldsymbol{e}_k}^{+k}\, p_{\boldsymbol{d}-\boldsymbol{e}_k} \quad \text{with} \quad \lambda_{\boldsymbol{d}} = r_{\boldsymbol{d}} - g(\boldsymbol{x}_{\boldsymbol{d}}) = r_{\boldsymbol{d}} - \sum_{k=1}^{3} u_{\boldsymbol{d}}^{+k}\,, \tag{5.55}$$

with eigenvalues $\lambda_{\boldsymbol{d}}$ as diagonal entries, because $\boldsymbol{H}$ is again a lower triangular matrix. In this case, the maximum principle can be written as

Theorem 6 (The maximum principle for the four-state model with unidirectional mutation rates). *Let the mutation rates be as in equation (5.51), and let* $\mathcal{S}_{\text{in}} := \{\boldsymbol{d} | p_{\boldsymbol{d}}(0) > 0\}$ *be the initially populated part of the mutational distance space. Then, the equilibrium population mean fitness is given by*

$$\bar{r} = \sup_{\boldsymbol{d} \in \bigcup_{\boldsymbol{d}_{\text{in}} \in \mathcal{S}_{\text{in}}} OC(\boldsymbol{d}_{\text{in}})} [r(\boldsymbol{x}_{\boldsymbol{d}}) - g(\boldsymbol{x}_{\boldsymbol{d}})]\,. \tag{5.56}$$

If, additionally, the maximum in (5.56) is taken at a unique mutational distance $\hat{\boldsymbol{d}}$*, it is furthermore*

$$\bar{r} = r(\hat{\boldsymbol{x}}) - g(\hat{\boldsymbol{x}})\,, \tag{5.57}$$

where $\hat{\boldsymbol{x}} = \boldsymbol{x}_{\hat{\boldsymbol{d}}}$.

Proof. Suppose again there is one $\boldsymbol{d}^*$ such that $p_{\boldsymbol{d}^*} > 0$, but $p_{\boldsymbol{d}^*} - \boldsymbol{e}_k = 0$ for all $k \in \{1, 2, 3\}$. According to equation (5.55), this can only happen if $\bar{r} = \lambda_{\boldsymbol{d}^*}$. An evaluation

of equation (5.55) for $\boldsymbol{d}^* - \boldsymbol{e}_k$ and for other sequences in the ancestral cone yields that no sequence in the ancestral cone can contribute to the population, i.e., $p_{\boldsymbol{d}} = 0$ for all $\boldsymbol{d} \in AC(\boldsymbol{d}^*)$.

If there was a sequence $\boldsymbol{d}^+$ in the offspring cone of $\boldsymbol{d}^*$ with $\lambda_{\boldsymbol{d}^+} > \lambda_{\boldsymbol{d}^*}$, $p_{\boldsymbol{d}^+}$ would be negative, thus this is impossible. Hence, $\lambda_{\boldsymbol{d}} \leq \lambda_{\boldsymbol{d}^*}$ for all $\boldsymbol{d} \in OC(\boldsymbol{d}^*)$, which corresponds to the first part of the maximum principle (5.56) as

$$\bar{r} = \max_{\boldsymbol{d} \in OC(\boldsymbol{d}^*)} [r(\boldsymbol{x}_{\boldsymbol{d}}) - g(\boldsymbol{x}_{\boldsymbol{d}})] \,. \tag{5.58}$$

If now $\lambda_{\boldsymbol{d}^+} = \lambda_{\boldsymbol{d}^*} = \bar{r}$, then $p_{\boldsymbol{d}} = 0$ for all $\boldsymbol{d} \in AC(\boldsymbol{d}^+)$, including $\boldsymbol{d}^*$, so that in this case the offspring cone $OC(\boldsymbol{d}^+)$ spans the population rather than $OC(\boldsymbol{d}^*)$. Evaluating the eigenvalue equation for the sequences in the offspring cone yields $p_{\boldsymbol{d}} > 0$ for all $\boldsymbol{d} \in OC(\boldsymbol{d}^*)$ if $\lambda_{\boldsymbol{d}^*} > \lambda_{\boldsymbol{d}^+}$, or for all $\boldsymbol{d} \in OC(\boldsymbol{d}^+)$ if $\lambda_{\boldsymbol{d}^*} = \lambda_{\boldsymbol{d}^+}$.

All sequences in the sibling domain $SD(\boldsymbol{d}^*)$ descend originally from the ancestors of $\boldsymbol{d}^*$ which are not present in the population. Thus, their frequencies must vanish, unless there is a sequence $\boldsymbol{d}^+ \in SD(\boldsymbol{d}^*)$ with $\lambda_{\boldsymbol{d}^+} = \lambda_{\boldsymbol{d}^*}$. In this case, if $p_{\boldsymbol{d}^+} > 0$, the sequences in both offspring cones have non-vanishing frequency, $p_{\boldsymbol{d}} > 0$ for all $\boldsymbol{d} \in OC(\boldsymbol{d}^*) \cup OC(\boldsymbol{d}^+)$; the frequencies of all other sequences vanish.

Due to the reducibility of the mutation matrix in the case of unidirectional mutation, the equilibrium population is not unique, but it is always the distribution assumed that has the highest possible population mean fitness $\bar{r}$, considering the given initial conditions. To find this equilibrium one has to take into account the offspring cones of all initially present sequences $\boldsymbol{d}_{\text{in}}$, and $\bar{r} = \sup_{\cup OC(\boldsymbol{d}_{\text{in}})} [r(\boldsymbol{x}_{\boldsymbol{d}}) - g(\boldsymbol{x}_{\boldsymbol{d}})]$, which is assumed for at least one sequence $\boldsymbol{d}^*$.

If this maximum is unique, the equilibrium population has non-zero entries only for the sequences in the offspring cone $OC(\boldsymbol{d}^*)$, whereas the relative reproductive success $\boldsymbol{z}$ corresponding to $\lambda_{\boldsymbol{d}^*}$ has non-zero entries only for the ancestors of $\boldsymbol{d}^*$, so that the only $\boldsymbol{d}$ with $a_{\boldsymbol{d}} \neq 0$ is $\boldsymbol{d}^*$, and hence $\boldsymbol{d}^* = \hat{\boldsymbol{d}}$ is the only ancestor, which again yields equation (5.32). □

An interesting case arises, however, if the maximum is not unique, but attained at two (or even more) sequences $\boldsymbol{d}^*$ and $\boldsymbol{d}^+$ in the subspace under consideration. In this degenerate case, the ancestral distribution cannot be obtained as easily as shown above, as the left and right eigenvectors have no non-zero overlap and thus it is not possible to normalise $\boldsymbol{z}$ such that $\sum_i a_i = 1$.

There are now two cases to distinguish. If the sequences lie in parent-offspring relation, i.e., $\boldsymbol{d}^+ \in OC(\boldsymbol{d}^*)$, $\boldsymbol{d}^*$ is the single ancestor, whereas the population is formed by the offspring cone of $\boldsymbol{d}^+$, i.e., $p_{\boldsymbol{d}} > 0$ if and only if $\boldsymbol{d} \in OC(\boldsymbol{d}^+)$. Note that equation (5.57) still holds, although in this special case the only ancestor $\boldsymbol{d}^*$ has zero frequency in the population.

If, however, $\boldsymbol{d}^*$ and $\boldsymbol{d}^+$ lie in sibling relation to each other, the population is formed by the union of their offspring cones $OC(\boldsymbol{d}^*) \cup OC(\boldsymbol{d}^+)$; and $\boldsymbol{d}^*$ and $\boldsymbol{d}^+$ both may have non-vanishing ancestral frequency, which are then determined by the initial conditions.

CHAPTER 6

Error thresholds and phase diagrams

The maximum principle (5.31) and (5.32) is a powerful tool to calculate the population mean fitness $\bar{r}$ in equilibrium for arbitrary fitness functions of the permutation-invariant or Hopfield type, for any range of mutation rates. Also, the ancestral mean genotype $\hat{\boldsymbol{x}}$ is available. The general method to identify $\bar{r}$ and $\hat{\boldsymbol{x}}$ is to consider the partial derivatives of $r - g$ with respect to the components $x_{v,k}$ of the mutational distance $\boldsymbol{x}$. A necessary condition for the function $r - g$ to have a maximum at a value $\boldsymbol{x}^*$ is that its derivatives at this $\boldsymbol{x}^*$ vanish,

$$\frac{\partial}{\partial x_{v,k}} [r(\boldsymbol{x}) - g(\boldsymbol{x})]_{\boldsymbol{x}=\boldsymbol{x}^*} = 0 \quad \forall v, k \,. \tag{6.1}$$

The global maximum of the function $r - g$ must lie on one of the points $\boldsymbol{x}^*$ that fulfil equation (6.1) or on the border of the mutational distance space. Thus by comparing the values of $r - g$ on these possible points, the global maximum can be identified.

Apart from the general possibility to investigate the dependence of the population mean fitness $\bar{r}$ on the mutation rate μ, this yields the opportunity to investigate the phenomenon of the *error threshold*, which has interested scientists ever since it was first conceived in [Eig71].

In the quasispecies model [Eig71], which was the first mutation–selection model using the sequence space as type space, the phenomenon occurs that the maximum sequence length that can be maintained in a population is limited by the mutation rate, or copying fidelity, i.e., the probability that at replication the nucleotide in question is copied accurately. At short sequence lengths and low mutation rates, or high copying fidelities, the population consists predominantly of sequences that are very similar to the wild-type; the population distribution is well localised in sequence space. However, the probabil-

ity to make an error-free copy of an original decreases exponentially with the sequence length, such that for a fixed mutation rate, longer sequences will accumulate a high number of mutations compared to the wild-type, leading ultimately to the delocalisation of the population in sequence space, thus the genetic information cannot be maintained in the population. Therefore, the phenomenon of the error threshold as described by Eigen can be expressed as a maximum possible sequence length that can be maintained at a given mutation rate; or alternatively, as a maximum possible mutation rate at a given sequence length.

This phenomenon has attracted considerable interest in the scientific community (e.g. [MS95, chapter 4.3], for a review see [BG00]), not the least because the original model was set up to explain the origin of life. In prebiotic evolution, the sequences were not part of a cell with all its machinery for reproduction, but rather copying "by chance", serving as templates for the synthesis of new sequences. Under these circumstances, the copying fidelity is very limited, and hence only populations of short sequences can bequeath the information they contain to subsequent generations. These sequences are in fact so short, that they cannot possibly contain the genetic information to code for proteins that are required to enhance the copying fidelity, which only would make longer sequences feasible and thus allow for the creation of the very proteins. Hence "catch 22 of the origin of life: No large genome without proteins, and no proteins without large genome" [May83].

One problem is that there is no generally accepted definition of an error threshold. The criterion used in the original quasispecies model [Eig71] is the disappearance of the wild-type from the population, which under the single peaked landscape goes in line with the complete delocalisation of the population in sequence space. However, these two effects do not necessarily coincide for other fitness landscapes.

A number of works have exploited the equivalence of the biological model with models from statistical mechanics. Leuthäusser [Leu86] has established the equivalence of a discrete time version of the quasispecies model with a two-dimensional classical Ising model which has been used in [Leu87, Tar92, FPS93, FP97, Pel02]; in [GGZ96, Gal97], the sys-

tem was mapped onto a model for elastic polymers; and in [BBW97], the equivalence of the para-muse model, used here, with a one-dimensional quantum spin chain has been identified.

The benefit of these equivalences is that they make available the well developed tools and methods from statistical mechanics in the biological context. For instance, in the framework of these models, the error thresholds have been identified with phase transitions in the analogous physical systems, a route that shall be followed here.

Adopted to the mutation–selection model used here, error thresholds are characterised by non-analytic behaviour of the population mean fitness $\bar{r}$ and the ancestral mean mutational distance $\hat{x}$, which acts as the order parameter, at a critical mutation rate μ_c. One can distinguish between first and second order error thresholds.

Definition (First and second order error threshold). A first order error threshold exists at a critical mutation rate μ_c, if the ancestral mean mutational distance as a function of the mutation rate $\hat{x}(\mu)$ shows a discontinuity at this μ_c, which is also reflected by a kink in the population mean fitness $\bar{r}(\mu)$.
A second order error threshold exists at at critical mutation rate μ_c, if the derivative of the ancestral mean mutational distance with respect to the mutation rate $\left[\frac{d\hat{x}}{d\mu}\right]_{\mu\to\mu_c}$ is discontinuous at this mutation rate μ_c.

In the examples shown later in this thesis, the second order error threshold always show an infinite derivative at the critical mutation rates. Note that, like phase transitions in physics, these definitions of the error thresholds apply in the strict sense only to a system with infinite sequence length ($N \to \infty$), for finite sequence lengths, the thresholds are smoothed out due to the lack of non-analyticities.

If error thresholds are considered as phase transitions, their existence not only hinges on the limit of an infinite sequence length. It is also required that even for mutation rates below the threshold the sequence space is explored fully. Because for infinite sequence length, the number of types becomes infinite too, $|\mathcal{S}| \to \infty$, this is only possible in the

limit of an infinite population size.

In [HRWB02], a finer classification of different error threshold phenomena was given. The first order error threshold they called "fitness threshold". Here, this term shall include also the second order error threshold, making all error thresholds as defined above fitness thresholds. Furthermore, the concept of the "degradation threshold" was introduced.

Definition (Degradation threshold). A degradation threshold is an error threshold of first or second order, where the population distribution beyond the critical mutation rate μ_c is given by the equidistribution in sequence space $\mathfrak{S}$.

Thus here the degradation threshold is a special case of a fitness threshold, going in line with the total delocalisation of the population in sequence space. Note that in the limit of infinite sequence length ($N \to \infty$), for which the error threshold definitions apply exactly, this equidistribution is reached immediately above μ_c, and beyond the threshold the population is insensitive to any further increase in mutation rates. In the case of finite sequence lengths, where the thresholds are smoothed out, the equidistribution is of course only reached asymptotically.

One reason that this characterisation of error thresholds as phase transitions has not found a wider acceptance might be that the order parameter, which in the models of physics is the magnetisation, was lacking a biological interpretation until recently, due to subtle differences in the normalisation [BBW98]. For instance, for the one-dimensional quantum chain [BBW97], which is the equivalent model of the para-muse model used here, the order parameter is the quantum mechanical magnetisation, which obeys an L_2-normalisation, whereas the natural observables in the mutation–selection model are the population averages, which are normalised in L_1. Only recently [HRWB02], the L_2-normalisation has found an interpretation as the ancestral mean and thus the order parameter has been identified with the ancestral mean mutational distance $\hat{x}$.

The original error threshold was observed for the single peaked fitness landscape, where a single sequence is attributed a high fitness value, all other sequences are equally disad-

vantageous. This is clearly an oversimplification and should not be regarded as anything but a toy model. A large amount of the literature concerning the error threshold phenomenon is very closely related to the model where it was originally reported [Eig71] and thus uses the single peaked landscape. This literature (up to 1989) has been reviewed in [EMS89]. Other fitness landscapes that have been investigated comprise, in the permutation-invariant case, linear and quadratic fitness functions, general functions showing epistasis, and as examples lacking permutation-invariance the Onsager landscape [BBW97, BW01], which has nearest neighbour interactions within the sequence, as well as various spin glass landscapes like the Hopfield landscape [Leu87, Tar92], the Sherrington-Kirkpatrick spin glass [BS93], the NK spin glass [CAW02], and the random energy model [FPS93, FP97], assigning random fitness values to each sequence.

One fitness landscape where an analytical solution can be obtained is the linear fitness, cf. [Rum87, Hig94, BBW97]. Note that this corresponds to a multiplicative landscape in a set-up using discrete time. For a linear fitness function, there is no error threshold, but the population changes smoothly from localised to delocalised with an increasing mutation rate.

For quadratic fitness functions, error thresholds only exist for antagonistic epistasis; they are absent for quadratic fitness functions with synergistic epistasis [BBW97, HWB01, GG04a]. These results go in line with those for general epistatic fitness functions [Wie97]. Studies using non-permutation-invariant fitness functions generally report the presence of error thresholds.

Like phase transitions in physics, error thresholds occur in the strict sense of non-analyticity of the observables only in the thermodynamic limit, which here corresponds to the limit of infinite sequence length $N \to \infty$. Hence with the maximum principle, which is exact in that very limit, they can be detected. However, for finite sequences, the phenomenon does not vanish altogether, but the jumps or kinks in the population mean fitness $\bar{r}$ and the ancestral mean genotype $\hat{\boldsymbol{x}}$ are smoothed out. Also the critical mutation rates are shifted for finite sequence length somewhat to smaller mutation rates, as shall

be seen later.

The definition of the error threshold as phase transition not only hinges on the assumption of an infinite sequence length, but as well on the deterministic nature of the model in question, which is granted by the assumption of an infinite population size, a case in which the sequence space is explored fully ($p_i > 0 \ \forall i$, according to Perron-Frobenius theory). However, there have been a number of studies investigating systems with finite population size, mainly employing simulation methods. [NS89] investigated the single peaked landscape and found similar results as in the infinite population case, only with critical mutation rates being shifted to lower values as compared to the infinite population case. [BS93] used complex correlated fitness landscapes such as the Sherrington-Kirkpatrick spin glass, whereas [CAW02] studied NK spin glass landscapes. In these works, three 'phases' are identified: for low mutation rates, the population is clustered around the maximum of the fitness landscape; for high mutation rates, the population is given by the equidistribution in sequence space. At intermediate mutation rates, however, the population drifts through sequence space exploring the secondary maxima of the fitness landscape without being localised. The existence of this phase hinges on the finite population size, as with an infinite population the whole sequence space is explored even for small (but finite) mutation rates and thus there can be no drifting through the space.

Further extensions of the original error threshold phenomenon comprise the investigation of diploid models [Hig94, WBS95], the incorporation of recombination in the model [BBN96], or the analysis of pleiotropy effects, i.e., the effect that one mutation might influence a number of traits. Also, systems where there are more than one genotype in the fittest class have been investigated [HSF96, RFS01], coining the notion of the *phenotypic error threshold.* These models have been inspired by fitness landscapes for RNA molecules, where the fitness is determined by the secondary structure, and the mapping from sequence to structure is many-to-one.

Of course the discussion of the error threshold phenomenon is academic if the threshold is an artefact of the model rather than a real biological phenomenon. This issue has been

subject to numerous debates, especially because it has first been predicted by a model using the over-simplistic single peaked landscape. However, over the years biologists have accumulated evidence that particularly RNA viruses naturally thrive at very high mutation rates [DH88, EB88] (of the order of 10^{-4} to 10^{-5} per base per replication [DES+96], corresponding to a genomic mutation rate of about 0.1 to 10 mutations per replication [DH97]), and a number of studies have reported that populations of RNA viruses only survive a moderate increase of their mutation rate, whereas if the mutation rate is increased further, the populations become extinct [HDdlTS90, LEK+99, SDLD00, CCA01], for reviews see [DES+96, DH97]. This corresponds to the population being pushed beyond the error threshold. It has been suggested to use the error threshold for antiviral therapies [Eig93], and in fact, recent experimental results indicate that this is the mechanism via which the broad-spectrum antiviral drug ribavirin works [CCA01].

This clearly warrants some further investigation of the error threshold phenomenon, which shall be done in the remainder of this chapter. For permutation-invariant fitness functions, section 6.1 investigates the four-state model, where results have been scarce [HWB01, GG04a]. As examples of fitness functions, quadratic fitness and truncation selection are considered. The phase structure is investigated and the results for infinite sequence length obtained via the maximum principle are compared with results for finite sequence lengths from direct calculations. For fitness functions of the Hopfield type, the investigations performed in section 6.2 are restricted to the two-state model in order to keep the number of variables limited. There, the investigations focus on quadratic Hopfield-type fitness functions in the cases of two and three patterns.

6.1 The four-state model with permutation-invariant fitness

Although the sequence space model has clearly been developed with the structure of DNA sequences in mind, which are written in a four-letter alphabet, the vast majority of applications to date has been concerned with two-state models. The few investigations of four-state models [HWB01, GG04a] have however revealed more complex behaviour in

this case, including the emergence of partially ordered phases that hinge on the multidimensional mutational distance used in the four-state model. The results for the four-state model with permutation-invariant fitness published in [GG04a] shall be presented here.

6.1.1 Choice of fitness functions

As mentioned above, error thresholds do not occur for all permutation-invariant fitness functions. For two-state models, a number of criteria for fitness functions to give rise to error thresholds has been given in [HRWB02]. For the four-state model, it is plausible to assume that the situation is similar.

It is easy to verify the absence of error thresholds for linear fitness functions in the four-state model. For error thresholds to exist, one needs at least the complexity of a quadratic function. In the following, two examples of fitness functions displaying error thresholds shall be investigated, namely a quadratic fitness function and a truncation selection, where all genotypes with less than a critical number of mutations are equally fit and all others equally unfit.

For simplicity, the mutation schemes used here are limited to the Kimura 2 parameter (K2P) and Jukes-Cantor (JC) mutation schemes as simplifications of the full Kimura 3ST mutation scheme, and the choice of fitness functions is restricted to those that are symmetric with respect to permutations of the mutational distances x_k, $k \in \{1, 2, 3\}$.

6.1.2 Quadratic symmetric fitness function

It can be shown that for quadratic fitness functions with positive epistasis, i.e., fitness functions with negative second derivatives, no phase transitions (or error thresholds) exist (cf. [HWB01]). Looking only at quadratic symmetric fitness functions with negative epistasis (or positive second derivative), a fairly general form is

$$r(\boldsymbol{x}) = c \sum_{k=1}^{3} x_k + \sum_{\substack{k,\ell=1 \\ k \leq \ell}}^{3} x_k x_\ell \,, \tag{6.2}$$

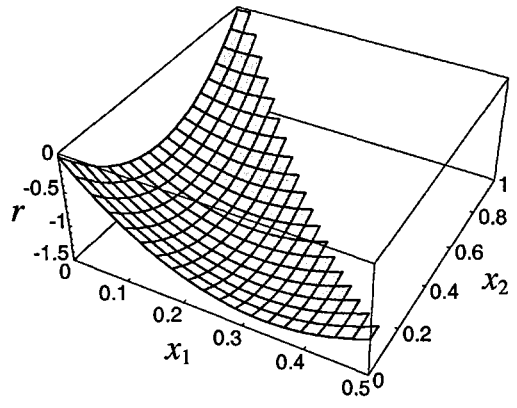

Figure 6.1: The quadratic symmetric fitness function (6.2) for $c = -1$ as a projection onto the relevant subspace with $x_1 = x_3$.

where the parameter c is used to tune the linear part relative to the quadratic term. The only possible generalisation within this restricted setup would be to give different coefficients for the pure quadratic terms x_k^2 and the mixed quadratic terms $x_k x_\ell$ with $k \neq \ell$. Here, the focus lies however on the fitness given in equation (6.2). Using one additional parameter to scale the overall quadratic term compared to the linear term would only result in a rescaling of the parameter c and the mutation rates. Note however that, because no care is taken here to obtain a realistic scaling of the fitness function, the absolute values of the mutation rates bear no meaning and should not be compared to any experimentally measured mutation rates. The ratio between the mutation rates μ/μ_2 is meaningful though.

In general, this fitness is symmetric with respect to permutation of the x_k, $k \in \{1, 2, 3\}$. For $c = -1$, the symmetry of the fitness function is even higher and includes the fraction of the wild-type-sites, $x_0 = 1 - \sum_{k=1}^{3} x_k$.

As the K2P mutation model has an inherent symmetry between x_1 and x_3, and between x_2 and x_0, respectively, and due to the symmetry of the fitness function (6.2) for arbitrary c, it is evident that the equilibrium solution for the ancestral mean mutational distance mirrors that symmetry with $\hat{x}_1 = \hat{x}_3$ (and in the case of $c = -1$, also $\hat{x}_2 = \hat{x}_0$). Figure 6.1 shows the quadratic symmetric fitness function (6.2) for $c = -1$ as a projection onto the relevant subspace where $x_1 = x_3$. The additional symmetry in the case $c = -1$ is reflected

in the fact that there are two equally high maxima at $(x_1, x_2) = (0, 0)$ and $(x_1, x_2) = (0, 1)$.

6.1.2.1 Phase diagrams

Phase diagrams are a convenient way to show the locations of phase transitions depending on the model parameters. Regions of the parameter space that are completely separated by phase transitions form distinct phases, in which the behaviour of the system differs.

The phase diagrams for the K2P mutation model with the quadratic symmetric fitness function (6.2) have been examined for the possible combinations of mutation rates and the parameter c. This has been done as described at the beginning of chapter 6 by considering the derivatives of the function $r - g$. Using the symmetry of the K2P mutation model, that is also inherent in the fitness function (6.2), we know that $\hat{x}_3 = \hat{x}_1$ and therefore the dimension of the equations can be reduced. With this, the fitness and mutational loss functions read explicitly

$$r(\boldsymbol{x}|x_3 = x_1) = c\,(2x_1 + x_2) + 3x_1^2 + 2x_1x_2 + x_2^2 \tag{6.3}$$

$$g(\boldsymbol{x}|x_3 = x_1) = 2\mu\,[1 - 2\sqrt{x_0x_1} - 2\sqrt{x_1x_2}] + \mu_2\,[1 - 2\sqrt{x_0x_2} - 2x_1] \tag{6.4}$$

with $x_0 = 1 - 2x_1 - x_2$. Thus as necessary conditions for a maximum we get

$$\begin{aligned}
&\frac{1}{2}\frac{\partial}{\partial x_1}(r - g) = \\
&c + 3x_1 + x_2 + \mu\left[\sqrt{\frac{x_0}{x_1}} - 2\sqrt{\frac{x_1}{x_0}} + \sqrt{\frac{x_2}{x_1}}\right] + \mu_2\left[-\sqrt{\frac{x_2}{x_0}} + 1\right] \quad \stackrel{!}{=} 0
\end{aligned} \tag{6.5}$$

$$\begin{aligned}
&\frac{\partial}{\partial x_2}(r - g) = \\
&c + 2x_1 + 2x_2 + 2\mu\left[\sqrt{\frac{x_1}{x_2}} - \sqrt{\frac{x_1}{x_0}}\right] + \mu_2\left[\sqrt{\frac{x_0}{x_2}} - \sqrt{\frac{x_2}{x_0}}\right] \quad \stackrel{!}{=} 0
\end{aligned} \tag{6.6}$$

Using the parametrisation $\mu = f\mu_2$, the difference between (6.5) and (6.6) yields

$$\mu_2 = \frac{(x_1 - x_2)\sqrt{x_1x_2}}{\sqrt{x_0x_1} - \sqrt{x_1x_2} + f(2x_1 - x_2 - \sqrt{x_0x_2})}\,. \tag{6.7}$$

Inserting this back into equation (6.6), and considering only the numerator, one gets

$$A\sqrt{x_0x_1} + B\sqrt{x_1x_2} = -f\,(C\sqrt{x_0x_2} + D) \tag{6.8}$$

with

$$A = -x_2\left[c + 2(x_1 + x_2)\right] \tag{6.9}$$

$$B = cx_0 + x_2(x_2 - x_1) + x_0(3x_1 + x_2) \tag{6.10}$$

$$C = 2cx_1 + 6x_1^2 - cx_2 - 2x_2^2 \tag{6.11}$$

$$D = -x_2\left\{cx_0 + 2\left[x_1(x_1 - x_2) + x_0(x_1 + x_2)\right]\right\}. \tag{6.12}$$

To eliminate the square roots in equation (6.8), we need to square twice, yielding

$$G^2 x_0 x_2 = J^2 \tag{6.13}$$

with

$$G = -2CDf^2 + 2ABx_1 \tag{6.14}$$

$$J = D^2 f^2 - A^2 x_0 x_1 + C^2 f^2 x_0 x_2 - B^2 x_1 x_2\,. \tag{6.15}$$

Equation (6.13) is a polynomial of order 10, and it can be solved , for instance using the algebraic computing package Mathematica with the function "NSolve", yielding x_2 for any input values of c, f and x_1. The obtained solutions have then to be checked whether they are a maximum of $r - g$. Using equation (6.7), the corresponding mutation rate can be determined.

The phase diagram for any fixed c is obtained by considering fixed ratios of $\mu/\mu_2 = f$, and finding the solutions of equation (6.13) for x_2, using x_1 as an input value (ranging from 0 to 1/2 in steps of typically 0.001). An example of the data obtained that way is shown in figure 6.2. Other Phase diagrams shown later on were obtained in a similar way.

For the case $c = -1$, the phase diagram is shown in Figure 6.3, cf. [HWB01], where a different notation, but the same fitness and mutation model are used. Here, three different phases can be identified:

- The AGCT phase. The population is essentially ordered, i.e., the population distribution is localised in sequence space.

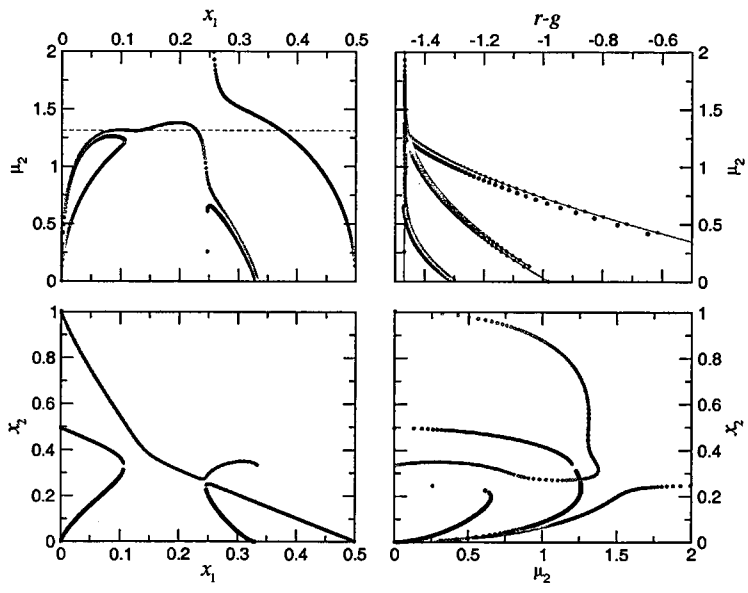

Figure 6.2: And example of the data used to generate the phase diagrams. Here, $c = -0.99$ and $f = 0.339$. The data consists of 4-tuples $(x_1, x_2, \mu_2, r-g)$, which are shown in various combinations. As the global maximum of $r-g$ is the desired solution, this is given by right most branch in the top right graph. For low mutation rates the solution is given by the green branch (which corresponds to small x_1 and large x_2), whereas for high mutation rates the solution is given by the red branch. The error threshold occurs where these two branches cross (indicated by the dashed line in the top left graph).

- The disordered phase. The population is completely random, the population distribution is the equidistribution in sequence space, the ancestral mean mutational distance is given by $\hat{x}_0 = \hat{x}_1 = \hat{x}_2 = \hat{x}_3 = 1/4$.

- The PP phase. A partially ordered phase, which only differentiates between purines and pyrimidines. In the ancestral distribution, there are two peaks that are equidistributions with respect to x_1, x_3– and x_0, x_2–direction, respectively, but localised in the other directions. This phase only exists in the case $c = -1$.

For large ratios μ/μ_2, the phase transition is a first order transition from the **AGCT** phase into the disordered phase, and thus it can be classified both as a fitness and a degradation threshold. For smaller ratios μ/μ_2, where the transition is from the **AGCT** to the PP phase, this is only a fitness threshold, of first or second order, depending on the exact ratio μ/μ_2, whereas the threshold from the PP into the disordered phase is a second order fitness and degradation threshold.

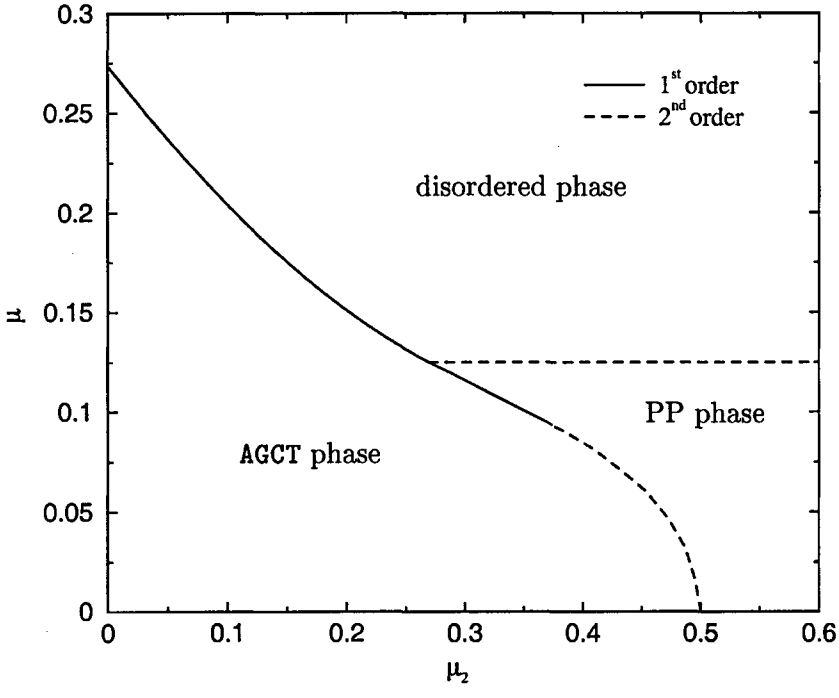

Figure 6.3: Phase diagram for the K2P mutation model and the quadratic symmetric fitness function (6.2) with $c = -1$. At the first order phase transition (solid line) the ancestral mean mutational distance $\hat{x}$ jumps and the population mean fitness $\bar{r}$ has a kink, whereas at the second order phase transitions (dashed lines) this kink in $\bar{r}$ is smoothed out and the jump in $\hat{x}$ is reduced to an infinite derivative at the critical mutation rate.

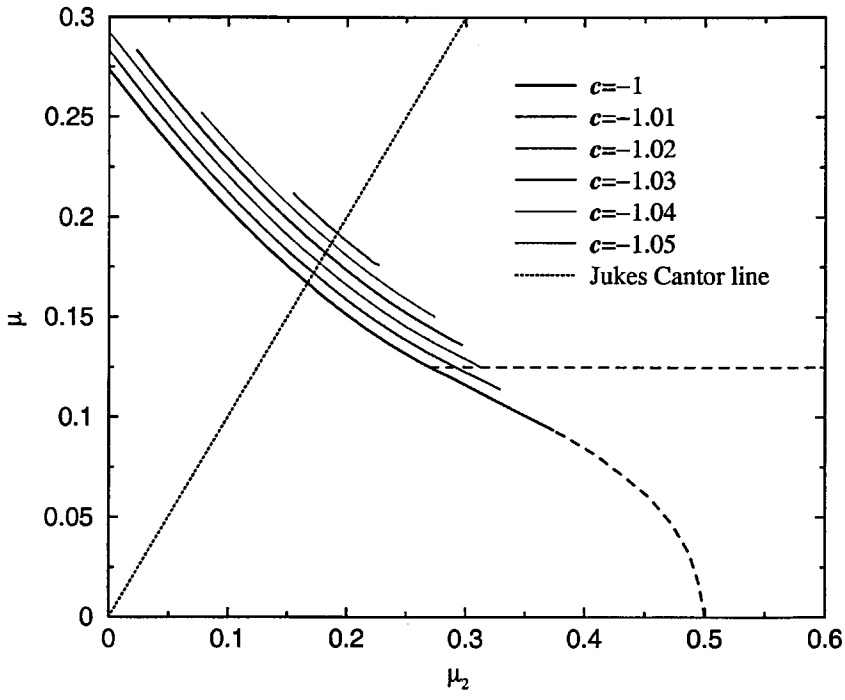

Figure 6.4: Phase diagram for the K2P mutation model and the quadratic symmetric fitness function (6.2) with $c \leq -1$.

If the symmetry between x_0 and x_2 is broken, i. e., $c \neq -1$, the PP phase disappears, see the phase diagrams shown in figures 6.4 and 6.5 with fitness function (6.2) for $c \leq -1$ and $c \geq -1$, respectively.

For $c < -1$, cf. Figure 6.4, the second order phase transitions disappear immediately with the broken symmetry, the first order phase transition line shrinks with decreasing c while shifting to slightly higher mutation rates and disappear for any $c \lesssim -1.06$. The disappearance of the second order transitions implies a disappearance of the disordered and the PP phase, leaving only an AGCT phase with varying degrees of order, and hence the remaining thresholds are only fitness thresholds, not degradation thresholds.

In the case $c > -1$, cf. Figure 6.5, the PP phase merges with the disordered phase, such that only the AGCT and disordered phases remain. The second order phase transition between AGCT and PP phase, that exists in the case $c = -1$, is transformed to a first order transition, in this case it is simultaneously a fitness and a degradation threshold. If $c \gtrsim -0.96$, even the AGCT and disordered phases merge as the phase transition line

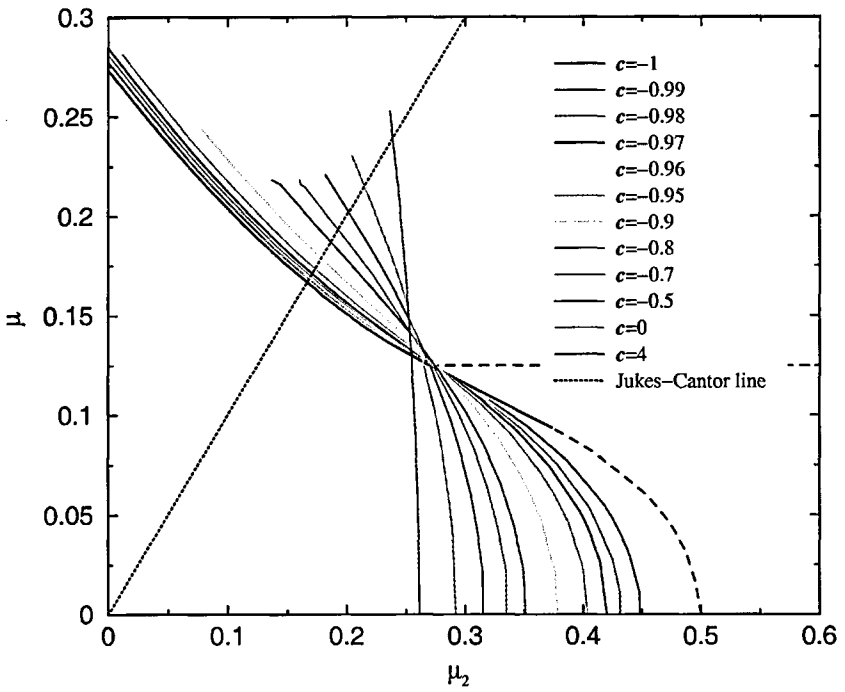

Figure 6.5: Phase diagram for the K2P mutation model and the quadratic symmetric fitness function (6.2) with $c \geq -1$.

does not reach to small values of μ_2, and thus the remaining phase transition lines can be characterised as fitness, but not as degradation thresholds.

6.1.2.2 Finite size effects

Using the maximum principle (equations (5.31) and (5.32)), only the population mean fitness $\bar{r}$ and the ancestral mutational distance $\hat{x}$ are accessible, which is sufficient to detect phase transitions. For small sequence lengths, it is however feasible to calculate $\bar{r}$ as largest eigenvalue and the population and ancestral distributions $\boldsymbol{p}$ and $\boldsymbol{a}$ through the corresponding eigenvector of $\widetilde{\boldsymbol{H}}$.

Figure 6.6 shows results obtained in this way for different finite sequence lengths using the quadratic fitness function (6.2) with parameter $c = -1$ compared with the results obtained via the maximum principle (5.31), which are exact for an infinite system. On the left, the JC mutation model was used, and consequently, the mean mutational distances in all three directions coincide. On the right, the K2P mutation model with $\mu = 0.3\mu_2$

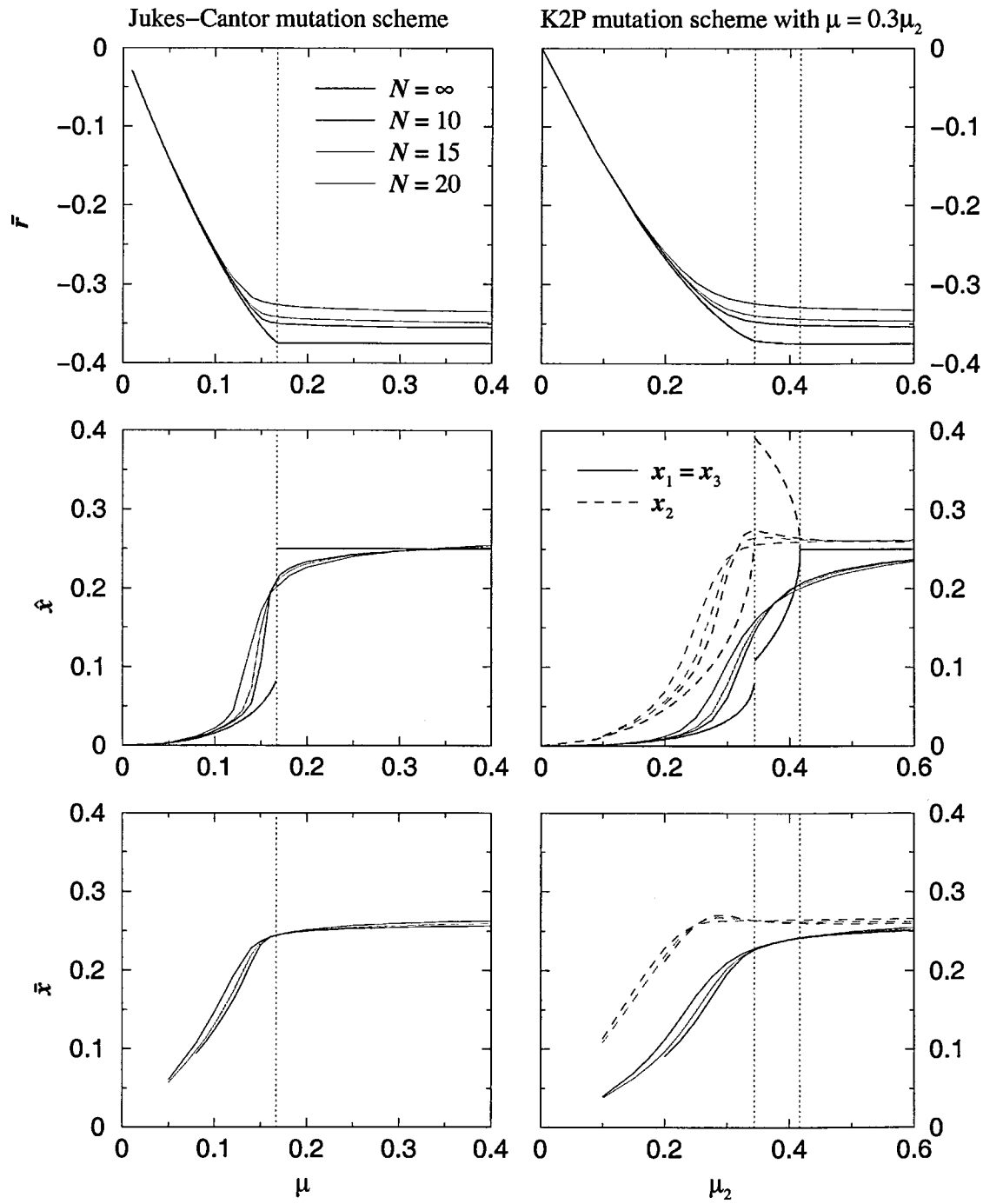

Figure 6.6: Comparison of population mean fitness $\bar{r}$, ancestral mean mutational distance $\hat{x}$ and population mean mutational distance $\bar{x}$ for different sequence lengths N and the infinite system ($N = \infty$) with fitness function (6.2) for $c = -1$. Left: JC mutation scheme, right: K2P mutation scheme with $\mu = 0.3\mu_2$. The locations of the phase transitions in the case of an infinite sequence length are marked as vertical dotted lines.

was used. Here, the mean mutational distances in x_1 and x_3 direction coincide, but differ from those in x_2 direction.

In Figure 6.6 one can clearly see how the phase transitions, which are sharp in the infinite system, are smoothed out for finite sequence lengths. Especially the second order phase transition in the K2P model cannot be detected at the sequence lengths considered here.

6.1.2.3 Distributions

The results in Figure 6.6 have been obtained by calculating the population and ancestral distributions explicitly, again using Mathematica. Examples of these are visualised in Figures 6.7 to 6.9. Although for the mean mutational distance $\hat{x}_1 = \hat{x}_3$, and therefore one can reduce the sequence space to the 2-dimensional subspace, sequences with $x_1 \neq x_3$ do occur in the equilibrium distributions with non-zero frequency, and thus the full three-dimensional mutational distance space $\mathcal{S}$ is needed for the visualisation. In Figures 6.7 to 6.9, the frequency of each point $\boldsymbol{d}$ is visualised as a cube of proportional size. For easier recognition, each cube corresponds to a block of 8 data points. The colours of the cubes indicate their position along the x_2 direction.

Figure 6.7 shows the ancestral distribution of the JC mutation model with the quadratic symmetric fitness function (6.2) with parameter $c = -1$. The phase transition, which happens around a mutation rate of $\mu \approx 0.16$, is clearly visible. For lower mutation rates, the population is in the AGCT phase and the ancestral distribution is localised. For mutation rates above the threshold, the population is in the disordered phase and the ancestral distribution is the equidistribution in sequence space. Note that the term equidistribution refers to the 4^N-dimensional sequence space $\mathfrak{S}$, and thus the ancestral distribution is a multinomial distribution in the depicted mutational distance space $\mathcal{S}$.

In Figure 6.8, the population distributions for the same model and parameter values as in Figure 6.7 are shown. The population distribution goes through the same stages as the ancestral distribution, but at lower mutation rates, and in contrast to the ancestral

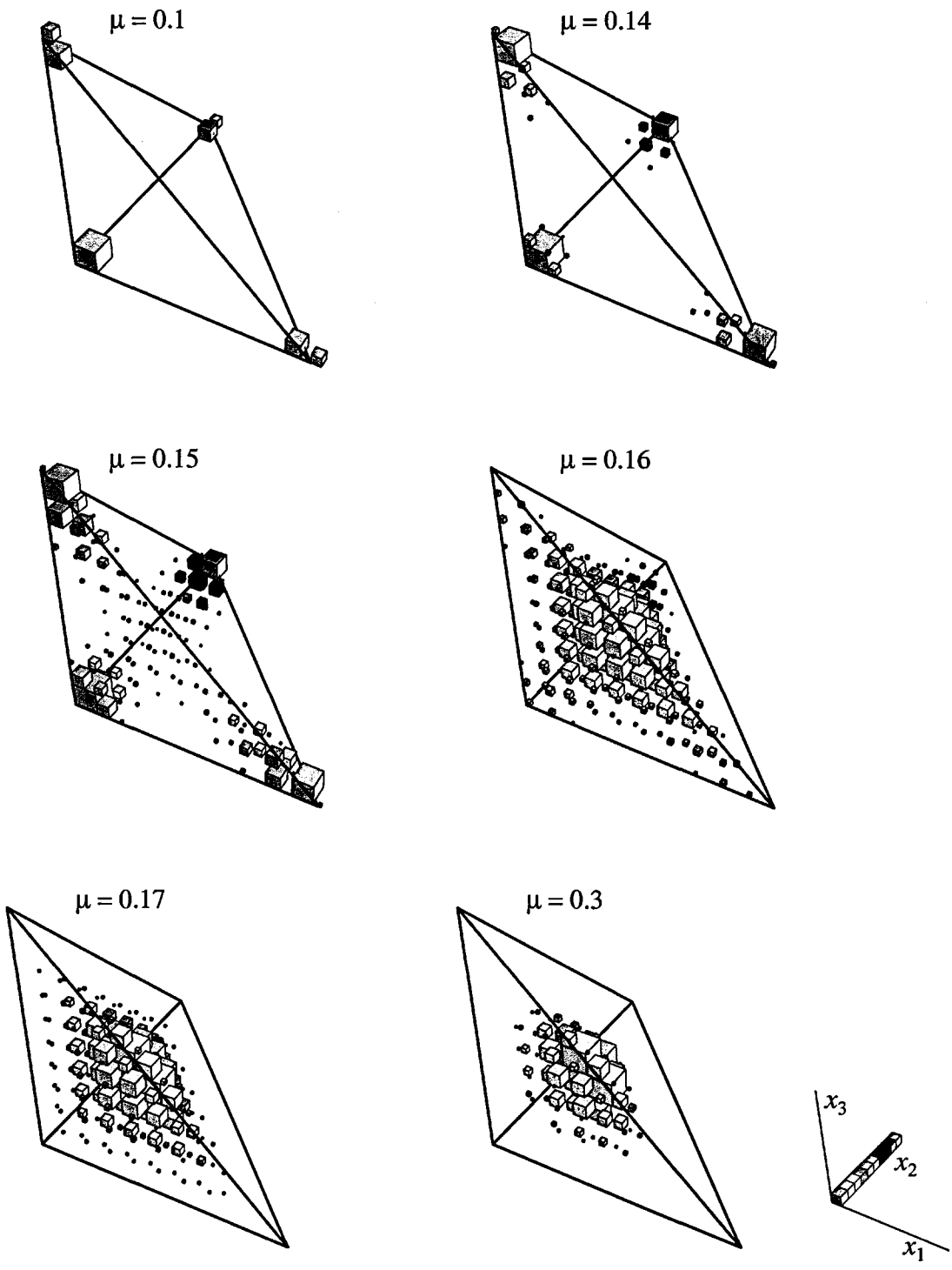

Figure 6.7: Ancestral distribution $\boldsymbol{a}$ for the JC mutation model with quadratic symmetric fitness (6.2) and $c = -1$ for selected mutation rates and a sequence length of $N = 20$. The frequency of each type $\boldsymbol{d}$ is symbolised by a cube of proportional size. Colours indicate the position along the x_2 direction. For easier recognition, each cube corresponds to a block of 8 data points.

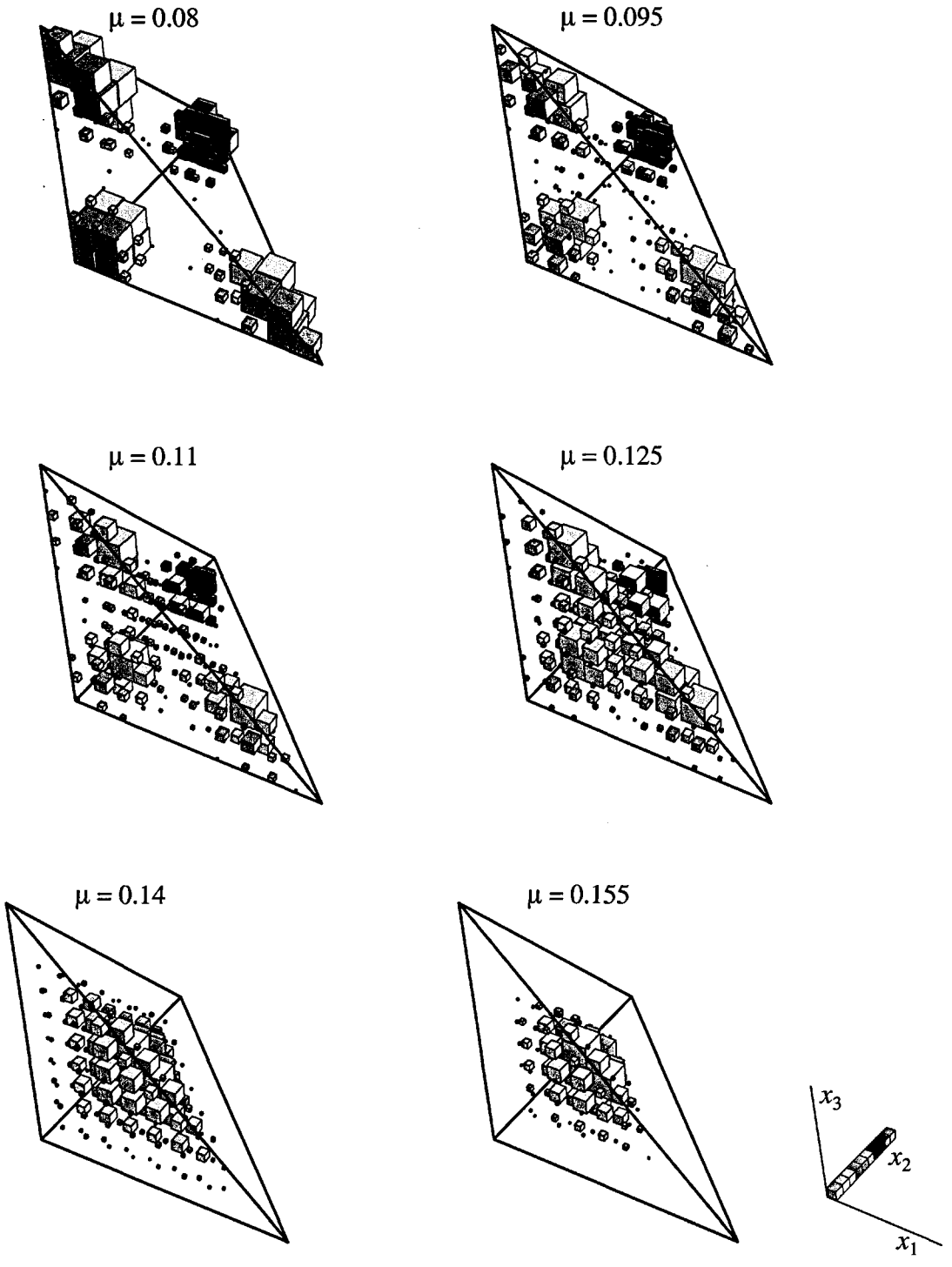

Figure 6.8: Population distribution $\boldsymbol{p}$ for the JC mutation model with fitness (6.2) and $c = -1$ for selected mutation rates and a sequence length of $N = 20$. For easier recognition, each cube corresponds to a block of 8 data points.

distribution, the transition is smooth, as can also be seen in Figure 6.6.

Figure 6.9 shows the ancestral distribution for the K2P mutation model with $\mu = 0.3\mu_2$ and fitness function (6.2) with $c = -1$. Here, the population starts in the AGCT phase (mutation rates $\mu_2 \lesssim 0.3$) and goes through the PP phase (at $\mu_2 \approx 0.325$), where the distribution consists of two "rods", which are each an equidistribution with respect to the x_2, x_0 or the x_3, x_1 direction, respectively, and localised with respect to the remaining directions. For high mutation rates, the equidistribution of the disordered phase ($\mu_2 \gtrsim 0.4$) can be found again.

6.1.3 Truncation selection

Another interesting fitness is the truncation selection, a case of extreme epistasis, both positive and negative. Here, it shall be used in the form

$$r(\boldsymbol{x}) = \begin{cases} 1 & \text{if } \sum_k x_k \leq 3x_c \\ 0 & \text{if } \sum_k x_k > 3x_c\,, \end{cases} \tag{6.16}$$

which again shows the symmetry between the x_k, $k \in \{1,2,3\}$, as demanded in section 6.1.1. Figure 6.10 shows a projection of the truncation selection (6.16) with $x_c = 0.1$ onto the subspace $x_1 = x_3$.

This is a generalisation of the single peaked landscape, as used in the original quasispecies model [Eig71]. The single peaked landscape corresponds to $x_c = 0$ in our setting, but it should only by used as a toy model. Truncation selection with non-zero x_c, however, is a standard model in biology, see e.g. [Kon88].

6.1.3.1 Phase diagram

Similarly to the case of the quadratic symmetric fitness, the phase diagram for the truncation selection (6.16) has been determined by using the method described at the beginning of chapter 6. However, the case of truncation selection is easier to treat.

As the derivative of a constant function is zero, to find the maximum of $r - g$ for a

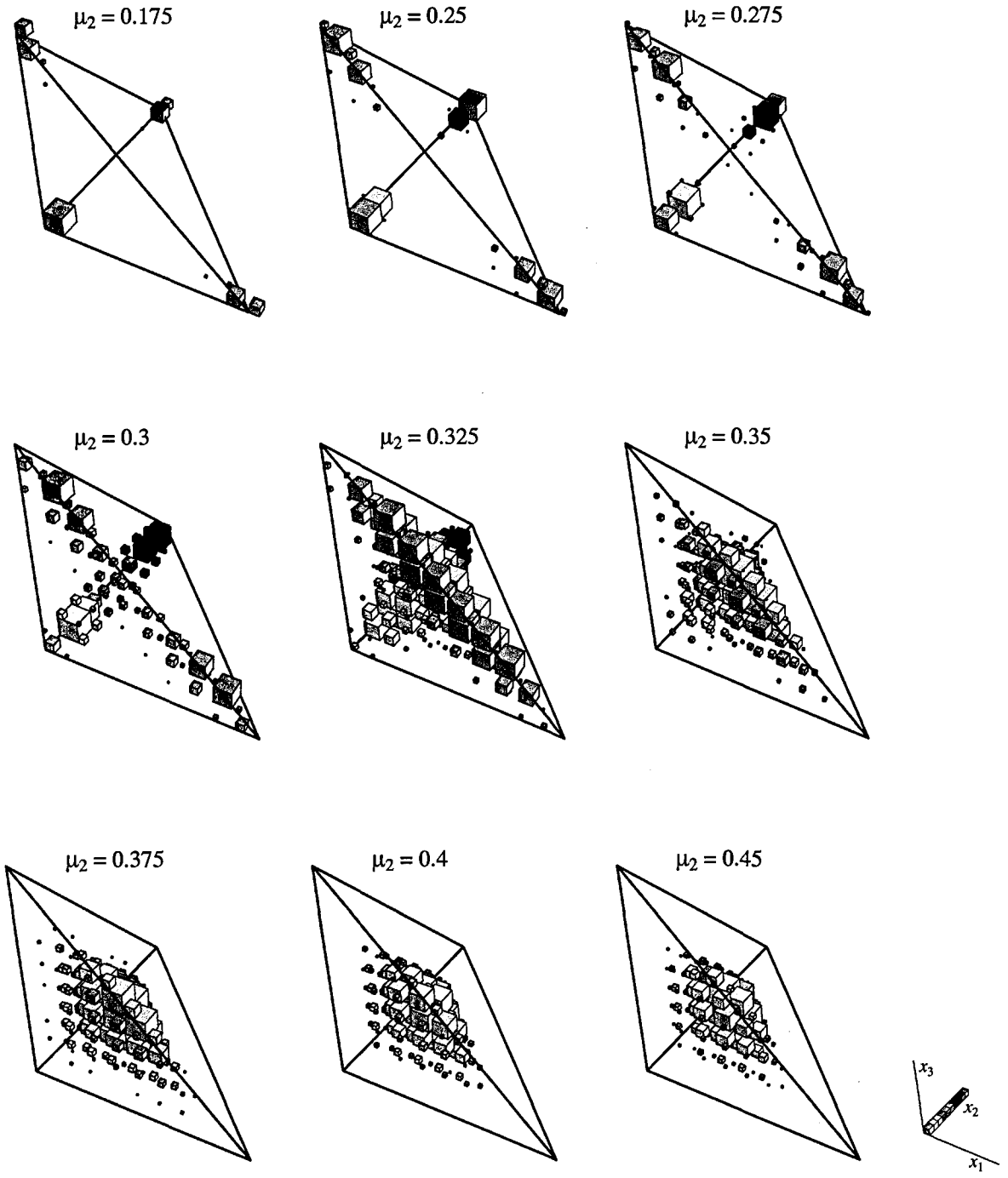

Figure 6.9: Ancestral distribution $\boldsymbol{a}$ for the K2P mutation model with $\mu = 0.3\mu_2$ with fitness (6.2) and $c = -1$ for selected mutation rates and a sequence length of $N = 20$. For easier recognition, each cube corresponds to a block of 8 data points.

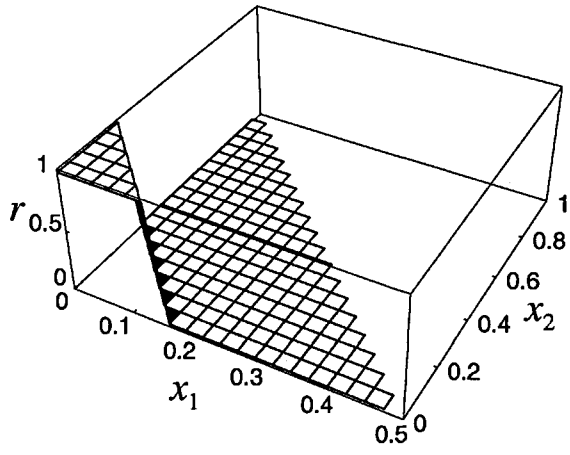

Figure 6.10: Truncation selection (6.16) for $x_c = 0.1$ as a projection onto the relevant subspace with $x_1 = x_3$.

truncation selection, which is constant over the sequence space apart from the truncation line, we only have to consider g, and that only at the mutation equilibrium, and on the truncation line.

Using K2P mutation model, g is given as in equation (6.4), and the mutation equilibrium, which is always assumed for high mutation rates, at least asymptotically, is given by $\boldsymbol{x}_{\text{eq}} = (x_{1,\text{eq}}, x_{2,\text{eq}}) = (1/4, 1/4)$ (implying that also $x_{0,\text{eq}} = x_{2,\text{eq}} = 1/4$).

If the mutation equilibrium lies in the high plateau of the fitness function, the system is in mutation equilibrium for any values of mutation rates. However, more interesting is the non-trivial case of the mutation equilibrium lying in the area with low fitness.

The global maximum of $r-g$ lies either at the mutation equilibrium, where $g(\boldsymbol{x}_{\text{eq}} = 0)$, or at the minimum of g along the truncation line, which is determined by $x_2 = 3x_c - 2x_1$. Along this line, $x_0 = 1 - 3x_c$ is constant. Thus the problem is reduced to finding the minimum x^* of a one-dimensional function $g(x_1)/\mu_2$ for given values of x_c and f, using again the parametrisation $\mu = f\mu_2$. In Mathematica, this can be done with the function "FindRoot".

Now we have to compare $(r-g)(x^*) = 1-g(x^*)$ with $(r-g)(x_{\text{eq}}) = r(x_{\text{eq}}) = 0$. For low mutation rates, the global maximum of $r-g$ always lies in x^*, whereas for large mutation

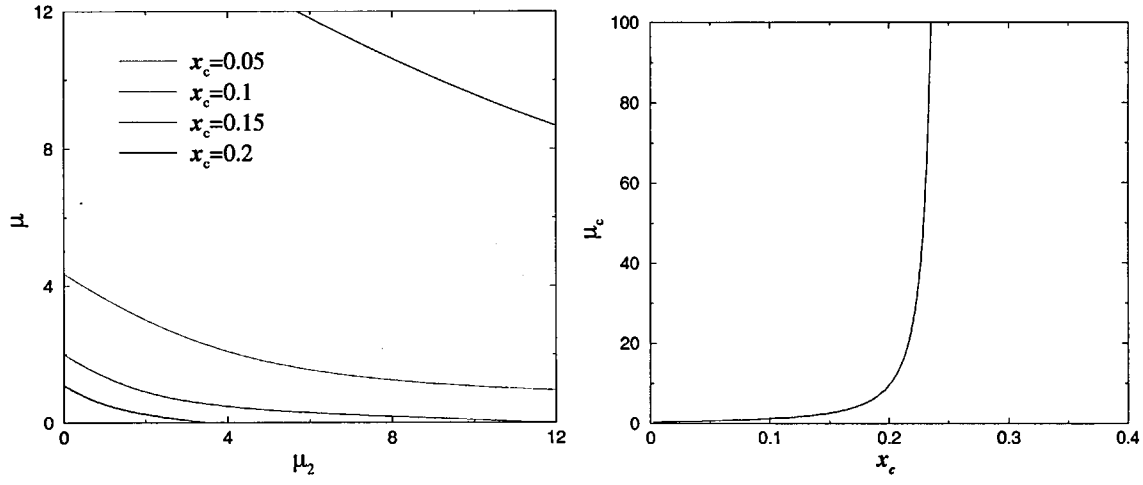

Figure 6.11: Left: Phase diagram for truncation selection (6.16) with different values of x_c. Right: Critical mutation rate μ_c against x_c for truncation selection (6.16) and the JC mutation model.

rates, it is in x_{eq}. As there is no continuous path from one to the other, inevitably there is a jump at a critical set of mutation rates. To find these mutation rates, we have to find the mutation rates for which $g(x^*) = 1$, or $g(x^*)/\mu_2 = 1/\mu_2$, which gives as critical mutation rate $\mu_{2,c} = \frac{1}{g(x^*)/\mu_2}$.

On the left, Figure 6.11 shows the phase diagram for truncation selection (6.16) with different values of x_c. Here, one finds error thresholds for all values of $x_c < 1/4$. For values of $x_c \geq 1/4$, the point of the mutation equilibrium x_{eq} is included in the high plateau and thus the population is in mutation equilibrium and at optimal fitness simultaneously for any mutation rate. On the right, Figure 6.11 shows the dependence of the critical mutation rate on x_c in the JC mutation model. At $x_c = 1/4$, the critical mutation rate diverges.

6.1.3.2 Finite size effects

Figure 6.12 compares the results for truncation selection obtained via the maximum principle (5.31) and (5.32) ($N = \infty$) with those obtained by calculating the largest eigenvalue and corresponding eigenvector of the symmetrised time-evolution operator $\widetilde{\boldsymbol{H}}$ for finite sequence lengths analogously to the data shown in Figure 6.6 for the quadratic symmetric fitness (6.2). On the left, one finds the data for the JC mutation scheme, whereas on the right, the results for the K2P model with $\mu = 0.3\mu_2$ are shown. Here, it is obvious

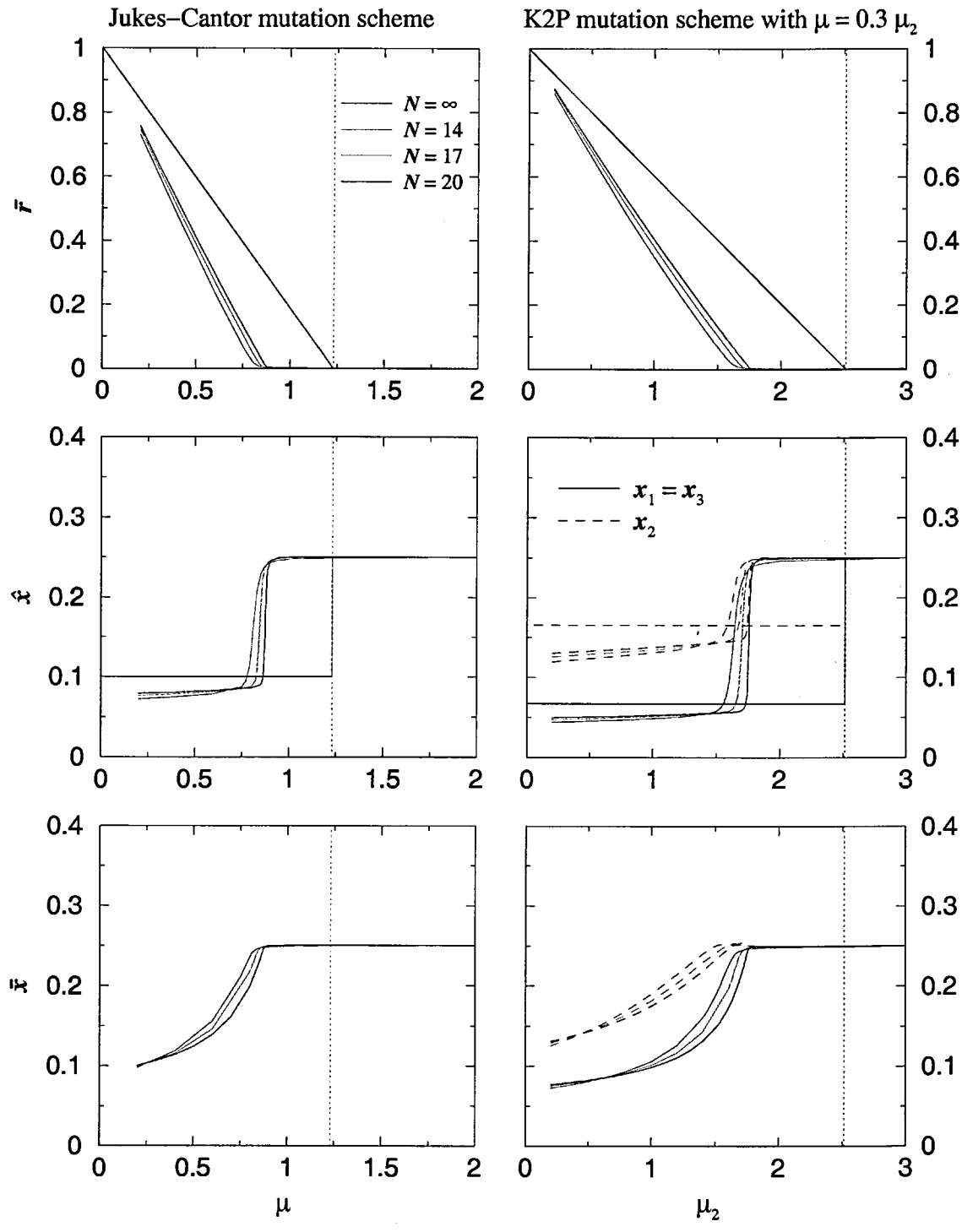

Figure 6.12: Comparison of population mean fitness $\bar{r}$, ancestral mean mutational distance $\hat{x}$ and population mean mutational distance $\bar{x}$ for different sequence lengths N and the infinite system ($N = \infty$) with truncation selection (6.16) for $x_c = 0.1$. Left: JC mutation scheme, right: K2P mutation scheme with $\mu = 0.3\mu_2$. The locations of the error thresholds in the case of an infinite sequence length are marked as vertical dotted lines.

that even for rather small sequence lengths, the error thresholds are very sharp. However, keeping in mind the discontinuity of the truncation selection, this is not too surprising. One can see that the location of the error threshold does depend on the sequence length: The smaller the sequence length N, the more the error threshold is shifted to lower mutation rates. So although the sequence lengths considered here are large enough to warrant a sharp error threshold, they are too small to predict the location of the error threshold in the infinite system.

6.1.3.3 Distributions

In the same way as for the quadratic symmetric fitness function (6.2), the ancestral and population distributions have been calculated for the truncation selection (6.16) and are shown in Figures 6.13 and 6.14. In contrast to Figures 6.7 to 6.9, which are the equivalent diagrams for the quadratic symmetric fitness function, every point in the sequence space is displayed as a separate cube, whose size is proportional to the fraction of individuals having that type in the ancestral and population distributions, respectively.

Figure 6.13 shows the ancestral distribution of a system with truncation selection (6.16) and $x_c = 0.1$ at a sequence length $N = 20$ for the JC mutation scheme for mutation rates close to the error threshold. For mutation rates below the threshold ($\mu \lesssim 0.86$), the ancestral distribution is an equidistribution in the "fit" part of the sequence space, where $\sum_k x_k \leq x_c$ and thus $r = 1$. For mutation rates above the threshold ($\mu \gtrsim 0.95$), the ancestral distribution is the equidistribution in the whole sequence space. The transition at the threshold is however interesting. Here, the distribution does not move smoothly from one equidistribution to the other, but in the intermediate states, the ancestral distribution is a superposition of the two equidistributions with an increasing proportion of the equidistribution on the whole sequence space.

Figure 6.14 shows the population distribution for a system with the same parameters as shown in Figure 6.13, i.e., JC mutation model and truncation selection (6.16) with $x_c = 0.1$ and sequence length $N = 20$. Here, the transition from a localised distribution for small

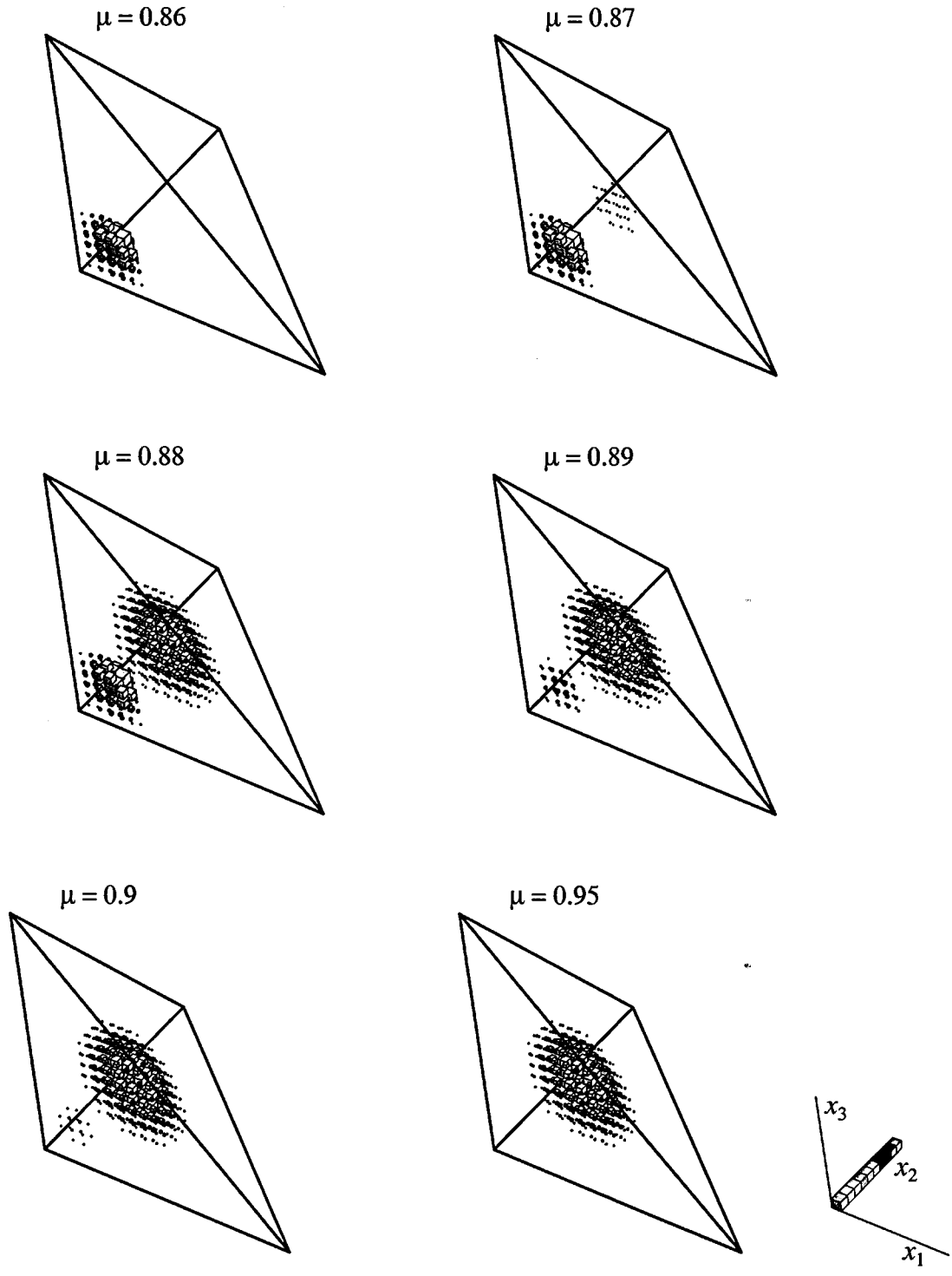

Figure 6.13: Ancestral distribution a for the JC mutation model with truncation selection (6.16) and $x_c = 0.1$ for mutation rates μ close to the error threshold and a sequence length of $N = 20$.

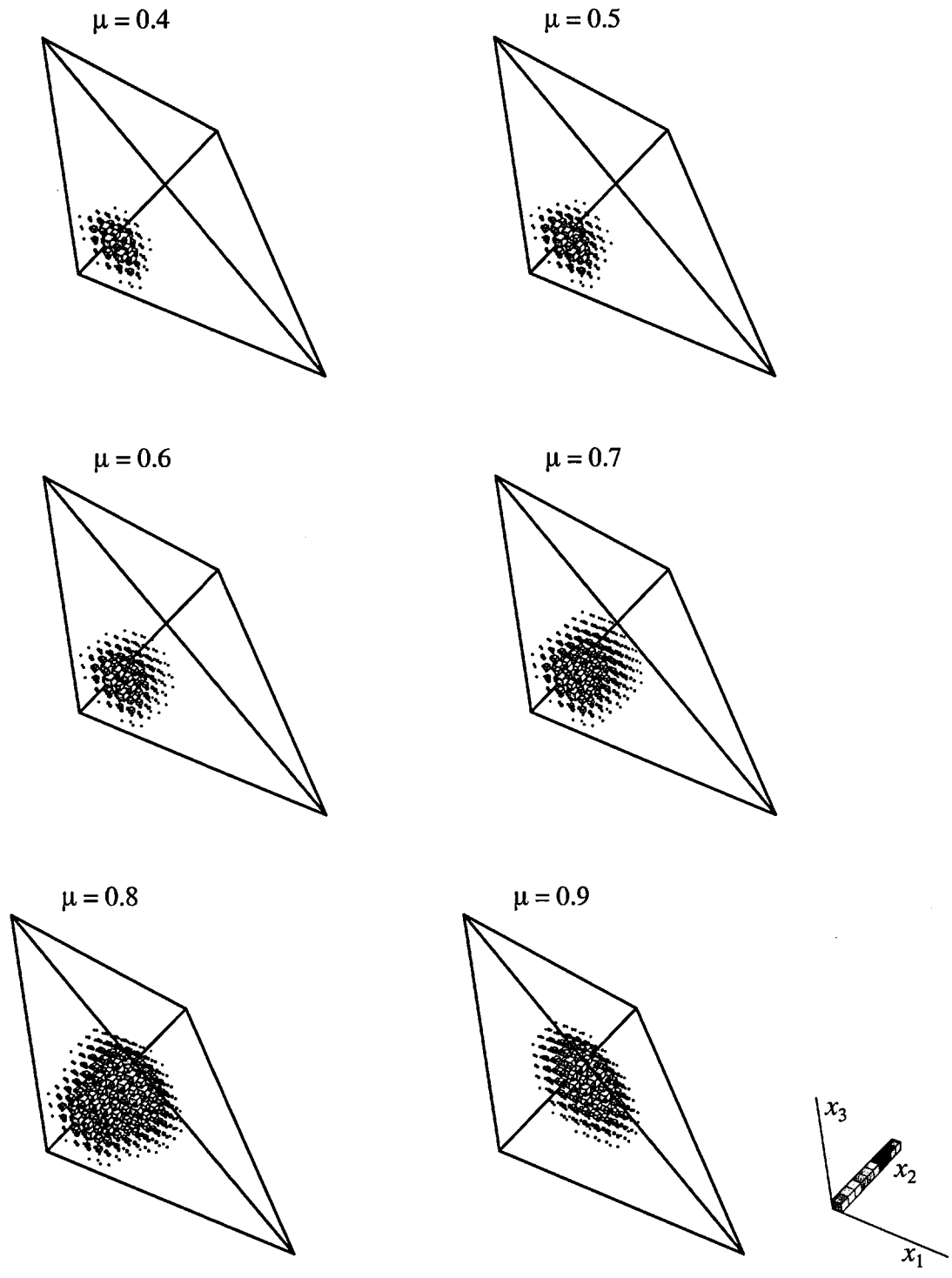

Figure 6.14: Population distribution $\boldsymbol{p}$ for the JC mutation model with truncation selection (6.16) and $x_c = 0.1$ for varying mutation rates μ and a sequence length of $N = 20$.

mutation rates to the equidistribution in sequence space looks far smoother (even apart from the far broader range of mutation rates over which it happens). The distribution as a whole shifts its centre until it reaches the equidistribution.

As the situation is very similar for general K2P mutation model (apart from a difference in mutation rates, cf. Figure 6.11, left), no distributions for a case with different mutation rates are shown.

6.1.4 Summary of the results for permutation-invariant fitness

For the four-state model with permutation-invariant fitness, two types of fitness functions that display error thresholds were investigated, namely a quadratic symmetric fitness function with negative epistasis and a truncation selection. The results obtained via the maximum principle, which are exact for infinite sequence length ($N = \infty$), were compared with results for (small) finite sequence lengths that were obtained by direct calculation of the leading eigenvalue and corresponding eigenvector.

For the quadratic fitness function, three different phases were identified, such as the ordered AGCT phase, the disordered phase and the partially ordered PP phase. Whereas the AGCT and disordered phases can also be found in two-state models, the existence of a partially ordered phase hinges on the multidimensional mutational distance that is used in the four-state model. However, it only occurs if the fitness function shows a high symmetry, which is not normally to be expected in natural populations. Yet it is not surprising to find a phase with a certain symmetry, like the PP phase that shows a symmetry between x_2 and x_0, only if the underlying mutation and selection models exhibit the same symmetry. For finite sequence lengths, the behaviour of the infinite system can already be recognised; however, the error thresholds are smoothed out.

For the truncation selection considered here, only the AGCT phase and the disordered phases occur, because this fitness function lacks the symmetry required for the PP phase. Here, the error thresholds for small finite sequence lengths are already very sharp, but their location does not coincide with the location predicted for infinite sequence length:

it is shifted toward lower mutation rates.

6.2 The two-state model with Hopfield-type fitness

The original Hopfield fitness [Hop82] is quadratic in the overlaps o^q with patterns $q = 0, \ldots, p$. In the original form, these overlaps vary from $-N$ to N with N meaning perfect match with the pattern and $-N$ meaning perfect match with the complementary sequence to the pattern in question, i.e., the sequence that differs at each site from the pattern. The original fitness is a pure quadratic function of the original overlaps o^q (4.15), $r \sim \sum_q (o^q)^2$. In the notation used here, the specific distances w^q to a pattern q, as given in equation (4.16), vary from 0 to N, with 0 as perfect match and N as complementary sequence. This means that the original Hopfield fitness in the notation used here reads

$$r \sim -\sum_{q=0}^{p} w^q + \sum_{q=0}^{p} (w^q)^2 \,, \tag{6.17}$$

apart from a constant, which does not influence the dynamics. Note the similarity to the permutation-invariant fitness (6.2) with higher symmetry ($c = -1$).

Most works that have studied a Hopfield-type fitness used the original Hopfield model [Leu87, Tar92], a generalisation was however treated in [Pel02], where a Hopfield-type truncation selection with two patterns was used.

6.2.1 Error thresholds and different Hopfield-type fitness functions

The Hopfield fitness has always been put forward as a realistic fitness landscape in the sense that it displays a reasonable and tunable amount of ruggedness, as well as correlations between closely related sequences. In the debate about the relevance of the error threshold phenomenon, it has been cited as an example for a complex fitness function, which displays the threshold behaviour. However, most applications have only been concerned with the original Hopfield fitness, rather than the generalised Hopfield-type fitness considered here. Thus it may be interesting and instructive to investigate the threshold behaviour of different kinds of Hopfield-type fitness functions.

To this end, consider first the shape of the mutational loss function for the two-state model with Hopfield-type fitness as given in equation (5.26). This is simply a superposition $g = \sum_v g_v$ of terms $g_v := \mu X_v \left(1 - 2\sqrt{x_v(1-x_v)}\right)$ in the partial distances x_v, hence these terms vary independently from each other. The X_v characterise the patterns as defined in section 5.2. The contributions g_v to the mutational loss functions have the same form as the mutational loss function in the two-state model with permutation-invariant fitness. If the fitness function is also decoupled with respect to the x_v, i.e., it can be written as a superposition of functions that depend only on one of the partial distances x_v, $r = \sum_v f_v(x_v)$, the threshold criteria obtained in [HRWB02] apply directly. This is however a strong constraint which can only be obtained in the Hopfield-type setting for fitness functions that depend linearly on the specific distances w^q, or, in the case of only two patterns, if the fitness depends only on the sum and difference of the specific distances w^0 and w^1. For linear Hopfield-type fitness there can be no thresholds, as follows from the threshold criteria in [HRWB02].

One step toward more complex fitness functions is to consider quadratic fitness functions. Here, the analysis shall be restricted to a symmetry with respect to the specific distances to the patterns $\boldsymbol{\xi}^q$,

$$r = c\sum_{q=0}^{p} y^q + d\sum_{q=0}^{p}(y^q)^2\,, \tag{6.18}$$

using the specific distances in normalised form $y^q := w^q/N$, in analogy to the symmetry with respect to the x_k, $k \in \{1,2,3\}$ in the permutation-invariant four-state case. Here, only the ratio c/d matters, as the absolute value only influences the scaling with respect to the mutation rates. Without loss of generality, d can thus be chosen as $d = \pm 1$. The original Hopfield fitness (6.17) is obtained in the case $d = 1$ and $c = -1$.

In the Hopfield-type fitness, epistasis can be defined with respect to each variable x_v as the sign of the second derivative of the fitness with respect to that variable: If $\frac{\partial^2 r}{\partial x_v^2} < 0$, epistasis with respect to x_v-direction is positive or synergistic; if $\frac{\partial^2 r}{\partial x_v^2} > 0$, epistasis is negative or antagonistic. In the quadratic symmetric Hopfield-type fitness from equation (6.18), $d = -1$ corresponds to positive epistasis for all x_v, whereas in the case $d = 1$,

epistasis is negative with respect to all x_v.

6.2.2 The case of two patterns

In the case of two patterns, $p = 1$, the first pattern can be chosen without loss of generality as $\boldsymbol{\xi}^0 = 00\ldots0$, such that there is only one pattern to be chosen, usually randomly. The matrix $\boldsymbol{\rho}$ containing the possible types of sites is then given by

$$\boldsymbol{\rho} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} , \tag{6.19}$$

and thus the index set of sites is partitioned into two subsets, $\Lambda = \Lambda_1 \cup \Lambda_2$, where Λ_1 contains all sites at which both patterns have entry 0, whereas Λ_2 corresponds to the sites where the two patterns have entries 0 and 1, respectively. The only quantities characterising the patterns are now the fractions of sites in each partition, $X_1 = |\Lambda_1|/N = N_1/N$ and $X_2 = N_2/N = 1 - X_1$. Thus the pattern can be characterised by a single parameter, X_1.

Each sequence is characterised with respect to the pattern by the partial Hamming distances to pattern $\boldsymbol{\xi}^0$ (in normalised form), x_1 and x_2. These vary from 0 (all entries 0 in Λ_v) to 1 (all entries 1 in Λ_v), completely independently from each other. The specific distances with respect to the patterns are linear combinations of the x_v and given in normalised form by

$$\begin{aligned} y^0 &= w^0/N = X_1 x_1 + X_2 x_2 , \\ y^1 &= w^1/N = X_1 x_1 + X_2(1 - x_2) . \end{aligned} \tag{6.20}$$

The Hopfield-type fitness is defined as an arbitrary function of these patterns, $r = f(y^0, y^1)$. Due to the small number of variables, for the case of two patterns, a lot can be done by analytical treatment.

For the quadratic symmetric Hopfield-type fitness (6.18) with positive epistasis ($d = -1$) and $c = 1$, there are no phase transitions, going in line with the results for permutation-invariant fitness functions. As an example for negative epistasis, consider

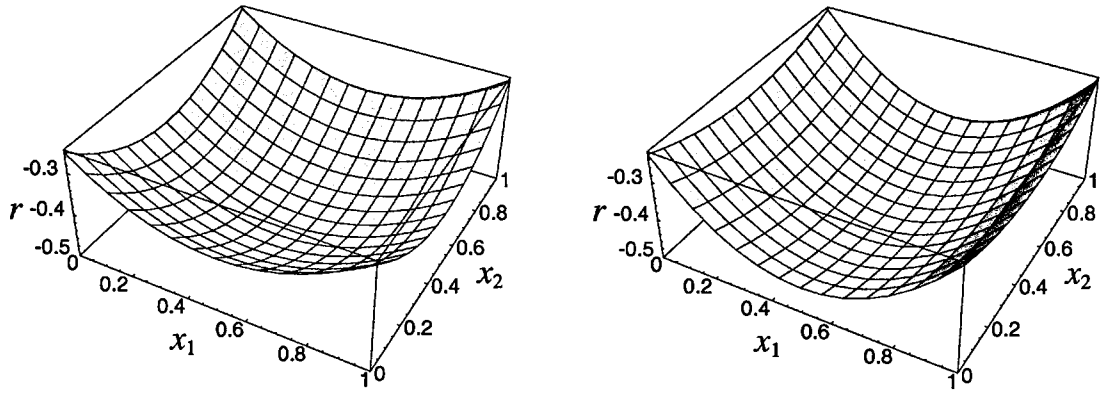

Figure 6.15: Original Hopfield fitness in the case of two patterns with $X_1 = X_2 = 1/2$ (left) and $X_1 = 1 - X_2 = 0.65$ (right).

first the original Hopfield fitness (6.17).

6.2.2.1 The original Hopfield fitness with two patterns

Since for two patterns there are only two variables, it is possible to visualise the fitness landscape in this case. Figure 6.15 shows the fitness landscape for the cases $X_1 = X_2 = 1/2$ and $X_1 = 1 - X_2 = 0.65$. In the corners of the mutational distance space $\mathcal{S}$, one can see the four maxima. The secondary maxima and saddle points that exist in the full sequence space $\mathfrak{S}$ do not show here, as the mapping from $\mathfrak{S}$ to $\mathcal{S}$ is done such that many matching ones are mapped onto the same mutational distances.

The ancestral mean partial distances $\hat{x}_v$, at which the maxima of $r - g$ are positioned, are obtained by considering the derivatives of $r - g$. They are given by

$$\hat{x}_v = \begin{cases} \frac{1}{2} \pm \frac{1}{2X_v}\sqrt{X_v^2 - \mu^2} & \text{for } \mu \leq X_v, \\ \frac{1}{2} & \text{for } \mu \geq X_v. \end{cases} \tag{6.21}$$

For the symmetric case with $X_1 = X_2 = 1/2$, the ancestral mean partial distances $\hat{x}_v$ are shown in figure 6.16 on the left, alongside the ancestral mean specific distances $\hat{y}^q = y^q(\hat{x}_v)$. For low mutation rates, there are two possible solutions for each of the ancestral mean partial distances $\hat{x}_v$, and as the maxima are degenerate, in equilibrium, the population will be centred equally around all of them. However, in the approach to equilibrium, the population might well be predominantly concentrated around one of them, depending on

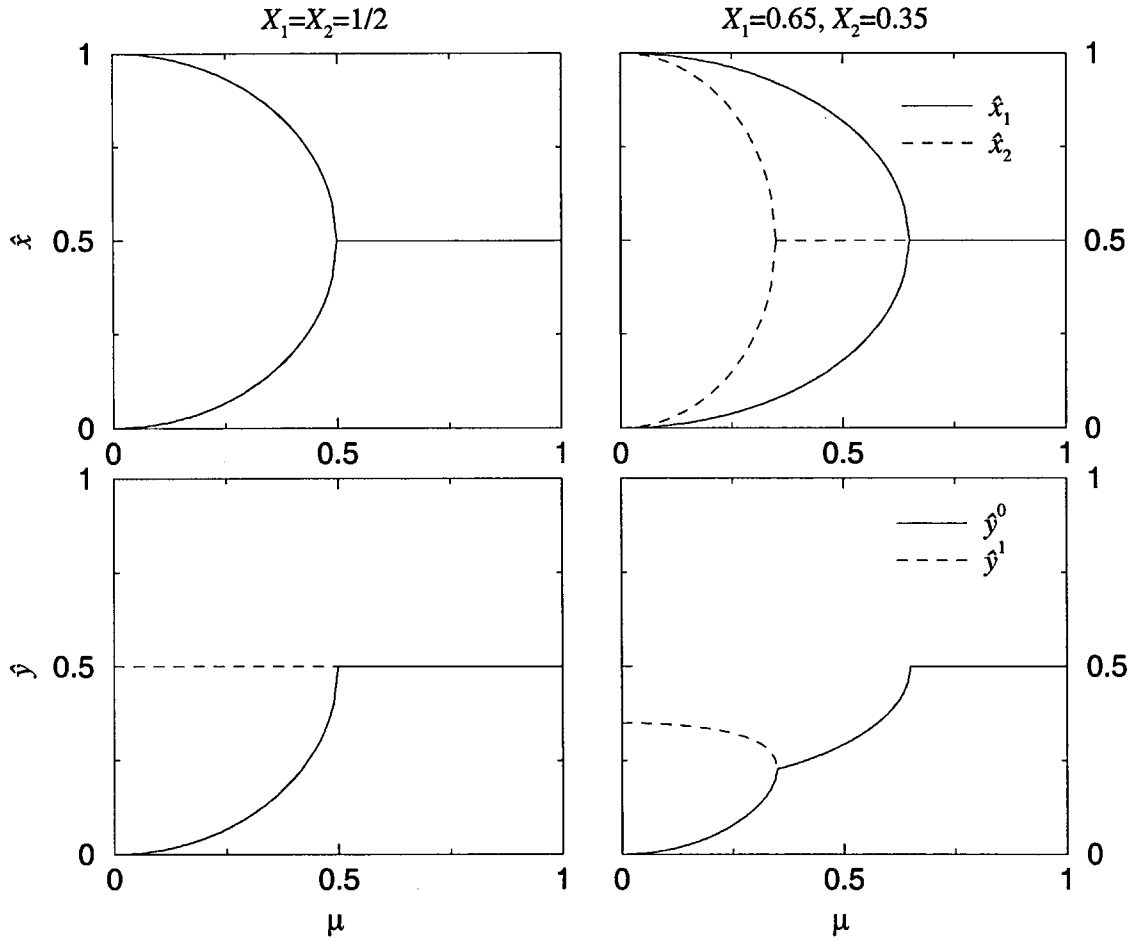

Figure 6.16: Ancestral mean partial distances $\hat{x}_v$ (top) and specific distances $\hat{y}^q$ (bottom) depending on mutation rates. The original Hopfield fitness (6.17) for two patterns has been used. Results correspond to a symmetric choice of patterns $X_1 = X_2 = 1/2$ (left) and an asymmetric choice with $X_1 = 1 - X_2 = 0.65$ (right).

initial conditions. For high mutation rates, the population is in the mutation equilibrium with $\hat{x}_1 = \hat{x}_2 = 1/2$ forming a disordered phase. The specific distances y^q that are shown correspond to the combination of $\hat{x}_1$ and $\hat{x}_2$, where both are given by the lower branch. Other combinations yield similar results. In the limit of low mutation rates, $\mu \to 0$, the population is always in the vicinity of one of the patterns (or its complement), such that one of the $\hat{y}^q \approx 0$ (1), which is completely random with respect to the other pattern, and thus the other $\hat{y}^q = 1/2$. This is the ordered phase. At the critical mutation rate $\mu_c = 1/2$, there is a second order phase transition between these two phases, which is a fitness as well as a degradation threshold, corresponding to the infinite derivative of both $\hat{x}_v$ at this mutation rate. As the specific distances y^q are simply superpositions of the partial distances x_v, the phase transitions are also visible in the y^q.

In the asymmetric case $X_1 \neq X_2$ (figure 6.16, right), two second order transitions can be identified. At $\mu = X_1$, $\hat{x}_1$ changes from the ordered phase into the disordered phase, whereas at $\mu = X_2$, $\hat{x}_2$ has its phase transition. The threshold occurring at the lower mutation rate is only a fitness threshold, whereas the one happening at the higher mutation rate is both a fitness and degradation threshold, leading to a totally random population. For $0 \leq \mu \leq \min(X_1, X_2)$, the population is in an ordered phase, for $\min(X_1, X_2) \leq \mu \leq \max(X_1, X_2)$, it is in a partially ordered phase, which is ordered with respect to one of the variables, but random with respect to the other. This phase is similar to the PP phase in the permutation-invariant four-state model. Finally for $\mu \geq \max(X_1, X_2)$, the population is the equidistribution in sequence space. Here again, for low mutation rates the population is close to one of the patterns, but due to the asymmetry in the chosen pattern, this leads to a non-random overlap with the other pattern. In the symmetric case with $X_1 = X_2 = 1/2$, the two error thresholds coincide, and the partially ordered phase vanishes.

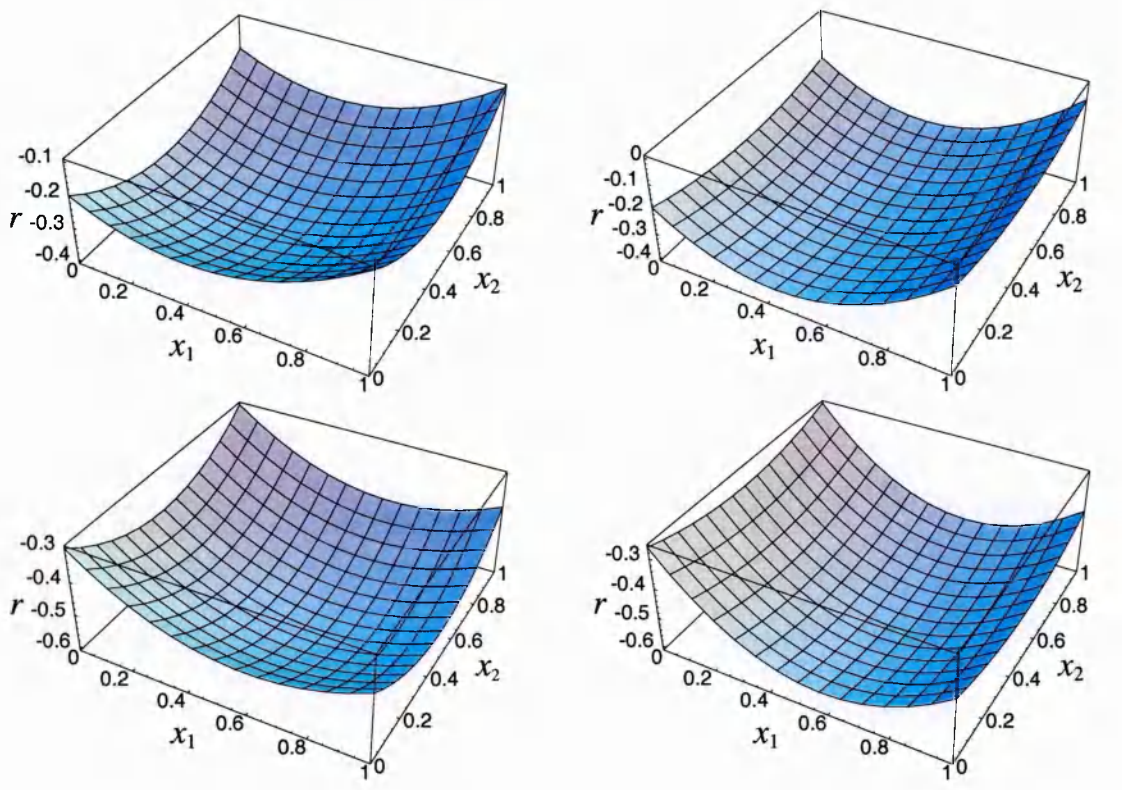

Figure 6.17: Quadratic Hopfield-type fitness functions (6.18) with negative epistasis and $c = -0.9$ (top) and $c = -1.1$ (bottom) for a symmetric pattern $X_1 = X_2 = 1/2$ (left) and an asymmetric pattern with $X_1 = 1 - X_2 = 0.65$ (right).

6.2.2.2 Deviations from the original Hopfield model

Now turn to the question how these phase transitions depend on the particular degeneracy of the fitness functions and consider the quadratic fitness function (6.18) with negative epistasis ($d = 1$) for values of $c \neq -1$.

Figure 6.17 shows the fitness landscapes for values of $c = -0.9$ and $c = -1.1$ for symmetric pattern $X_1 = X_2 = 1/2$ and an asymmetric pattern with $X_1 = 1 - X_2 = 0.65$. As the pictures indicate, the fitness functions (and thus the behaviour of the system) with the same $|c+1|$ are related by symmetry under $x_1 \to 1 - x_1$ (apart from a constant term, which does not influence the dynamics). Note that in x_2-direction, the fitness function is independent of c. This is because in the sum of the specific distances, the term with x_2 cancels out, which happens generally in the case of an even number of patterns (i.e., odd p) for $\binom{p}{(p+1)/2}$ different variables.

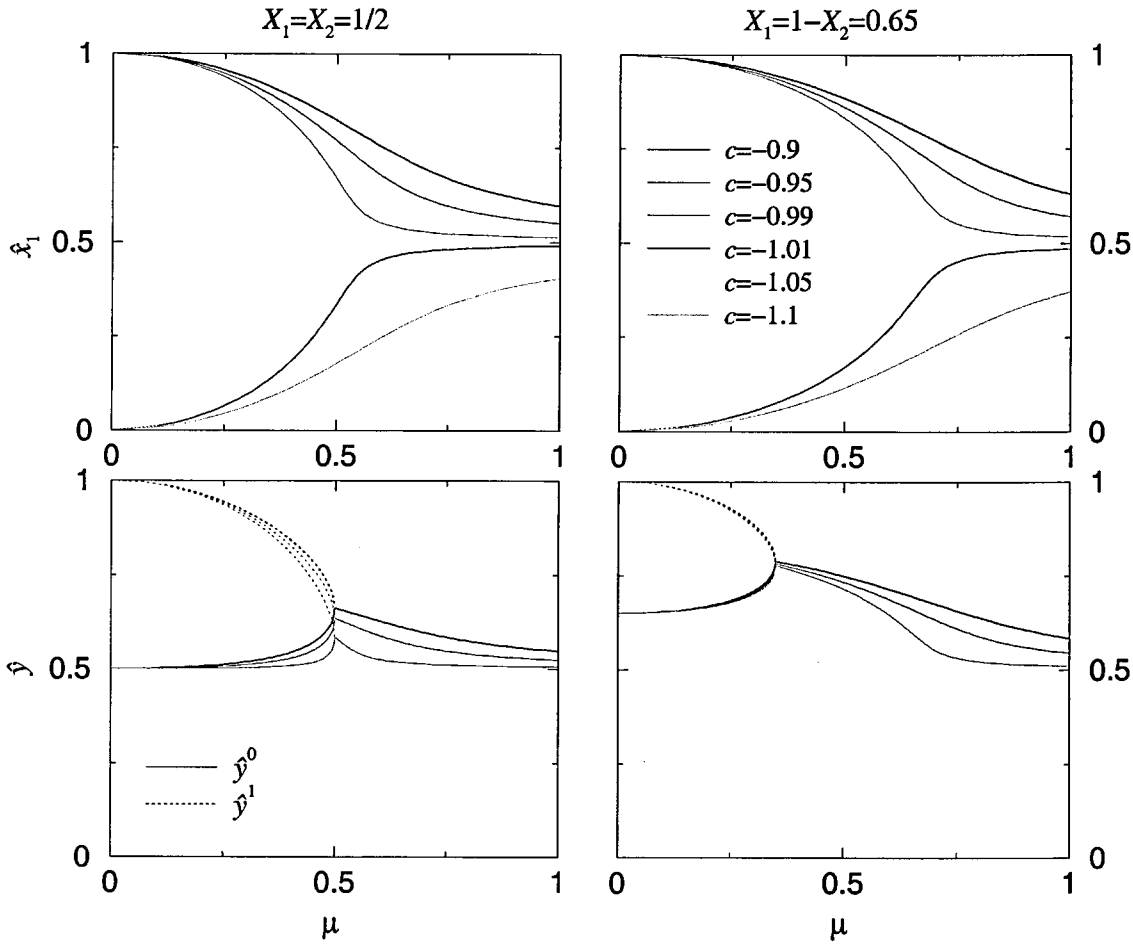

Figure 6.18: Ancestral mean partial distance $\hat{x}_1$ (top) and specific distances $\hat{y}^0$ (solid lines) and $\hat{y}^1$ (dashed lines, bottom) depending on mutation rates. The quadratic Hopfield-type fitness (6.18) with negative epistasis for two patterns and different values of c has been used. Results correspond to a symmetric choice of patterns $X_1 = X_2 = 1/2$ (left) and an asymmetric choice with $X_1 = 1 - X_2 = 0.65$ (right). For clarity, only specific distances $\hat{y}^q$ corresponding to $c > -1$ are shown.

Because in x_2-direction, the fitness is independent of c, the solution for the ancestral mean mutational distance $\hat{x}_2$ is identical with the solution for the original Hopfield fitness as given in equation (6.21). For x_1, the solution becomes more complicated, but the inverse function is simpler. It is given by

$$\mu = \frac{2\left[1 + c + X_1(2\hat{x}_1 - 1)\right]\sqrt{\hat{x}_1(1-\hat{x}_1)}}{2\hat{x}_1 - 1}. \tag{6.22}$$

So for all values of c, the phase transition with respect to x_2 happens at $\mu = X_2$. The dependence of $\hat{x}_1$ on the mutation rate is shown in figure 6.18 (top). For $c \neq -1$, the second order phase transition is smoothed out and thus vanishes. Note that the ambiguity

in the solutions that exists in the case $c = -1$ (cf. figure 6.16), does not exist here, due to the lacking degeneracy of the maxima of the fitness function at $x_1 = 0$ and $x_1 = 1$ (cf. figure 6.17). At the bottom, figure 6.18 shows the specific distances $\hat{y}^q$, using the lower branch of the solution for $\hat{x}_2$ (as shown in figure 6.16), which show the second order transition in $\hat{x}_2$, a fitness threshold. With this combination of solutions, for low mutation rates the population is centred around the sequence complementary to pattern $\boldsymbol{\xi}^1$. The general picture for the symmetric and asymmetric choice of patterns is very similar, apart from issues like the exact location of the thresholds.

6.2.3 The case of three patterns

For three patterns, the matrix $\boldsymbol{\rho}$ reads

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}, \tag{6.23}$$

thus there are four X_v describing the patterns $\boldsymbol{\xi}$, fulfilling $\sum_{v=1}^{4} X_v = 1$, and four variables $x_v \in [0, 1]$, describing each sequence. The specific distances with respect to pattern $\boldsymbol{\xi}^q$ are given by

$$y^0 = X_1 x_1 + X_2 x_2 + X_3 x_3 + X_4 x_4 \,, \tag{6.24}$$

$$y^1 = X_1 x_1 + X_2 x_2 + X_3 (1 - x_3) + X_4 (1 - x_4) \,, \tag{6.25}$$

$$y^2 = X_1 x_1 + X_2 (1 - x_2) + X_3 x_3 + X_4 (1 - x_4) \,. \tag{6.26}$$

Similarly to the case of two patterns, the original Hopfield fitness (6.17) shall be considered first, and then variations (6.18) with negative epistasis, but $c \neq -1$. The investigation is done by means of numerical calculations.

To this end, the derivatives of $r - g$ with respect to the different x_v are calculated. To eliminate the square roots, the equations are squared, such that instead of

$$\frac{\partial}{\partial x_v}(r - g) = 0 \,, \tag{6.27}$$

the equation

$$\frac{\partial}{\partial x_v}(r^2 - g^2) = 0 \tag{6.28}$$

is considered. This yields a system of four polynomials of 4th order, namely

$$4x_1(1-x_1)\left[3c/2 + 3X_1x_1 + X_2(1+x_2) + X_3(1+x3) + X_4(2-x_4)\right]^2 - \mu^2(1-2x_1)^2 = 0 \tag{6.29}$$

$$4x_2(1-x_2)\left[c/2 + X_1x_1 + X_2(3x_2-1) + X_3(1-x_3) + X_4x_4\right]^2 - \mu^2(1-2x_2)^2 = 0 \tag{6.30}$$

$$4x_3(1-x_3)\left[c/2 + X_1x_1 + X_2(1-x_2) + X_3(3x_3-1) + X_4x_4\right]^2 - \mu^2(1-2x_3)^2 = 0 \tag{6.31}$$

$$4x_4(1-x_4)\left[c/2 + X_1x_1 + X_2(1-x_2) + X_3(1-x_3) + X_4(2-3x_4)\right]^2 - \mu^2(1-2x_4)^2 = 0 \tag{6.32}$$

For any given values of c and (X_1, X_2, X_3, X_4), these can be evaluated using the Mathematica function "FindRoot", which finds the closest zero from a given starting configuration (x_1, x_2, x_3, x_4). Because the closest zero is not necessarily identical with the global maximum of $r - g$, the equations (6.29) to (6.32) have been solved for multiple starting values, and only the solutions with maximal $r - g$ have been retained. The starting values were spaced with a distance of $1/4$ in each of the variable, yielding $4^5 = 1024$ different starting values. Of course this does not guarantee that the obtained solution is the global maximum. However, the results for a number of parameter values were calculated with different smaller mesh sizes, i.e., different smaller spacings of the starting values, and the solutions were found to be independent of the mesh size. This indicates that already with the mesh size of $1/4$, the global maximum was found.

6.2.3.1 Variations of the pattern

If the patterns ξ^q for $q = 1, \ldots, p$ are chosen randomly (remember $\xi^0 = 00\ldots0$), in the limit of infinite sequence length, $N \to \infty$, the X_v are given by $X_v = 2^{-(p+1)} + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right)$ (cf.

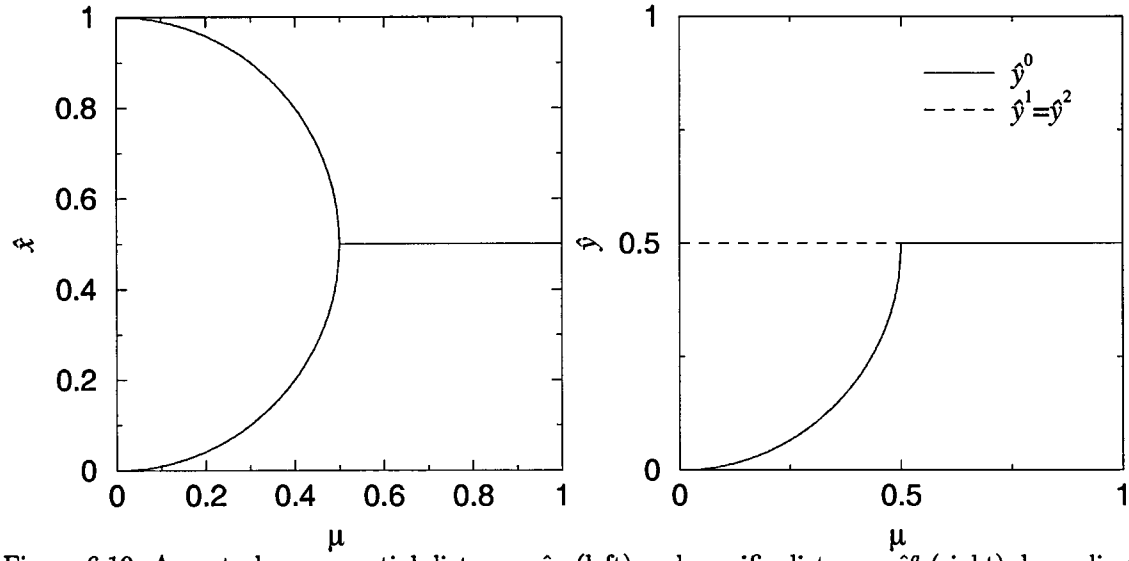

Figure 6.19: Ancestral mean partial distances $\hat{x}_v$ (left) and specific distances $\hat{y}^q$ (right) depending on mutation rates. The original Hopfield fitness (6.17) for three patterns has been used. Results correspond to patterns for infinite sequence length.

section 4.3, equation (4.42)). So to simulate the case of infinite sequence length, in which the maximum principle is exact, one has to assume $X_v = 2^{-(p+1)}$ for all $v = 1, \ldots, 2^{p+1}$. However, the maximum principle can also be used to investigate the case of finite sequence length by simulating the finite sequence length through choosing pattern distributions that do not follow exactly the infinite distribution $X_v = 2^{-(p+1)}$, but vary around this mean value with a variance according to the sequence length to be simulated. Practically, patterns corresponding to finite sequence length N have been obtained by choosing for the $3N$ sites entries 0 or 1 with probability 1/2 at each site, and counting the number of sites N_v in each class Λ_v, similarly to the example pattern given in equation (4.17).

As shall be seen in the following, the variations stemming from the simulated finite sequence length of the pattern do account for some interesting additional features. However, focus first on the case of a genuinely infinite sequence length with $X_v = 2^{-(p+1)}$ for all v.

6.2.3.2 The case of infinite sequence length

Original Hopfield fitness ($c = -1$). Figure 6.19 (left) shows the possible solutions of the ancestral mean partial distances x_v in the case of three patterns with the original

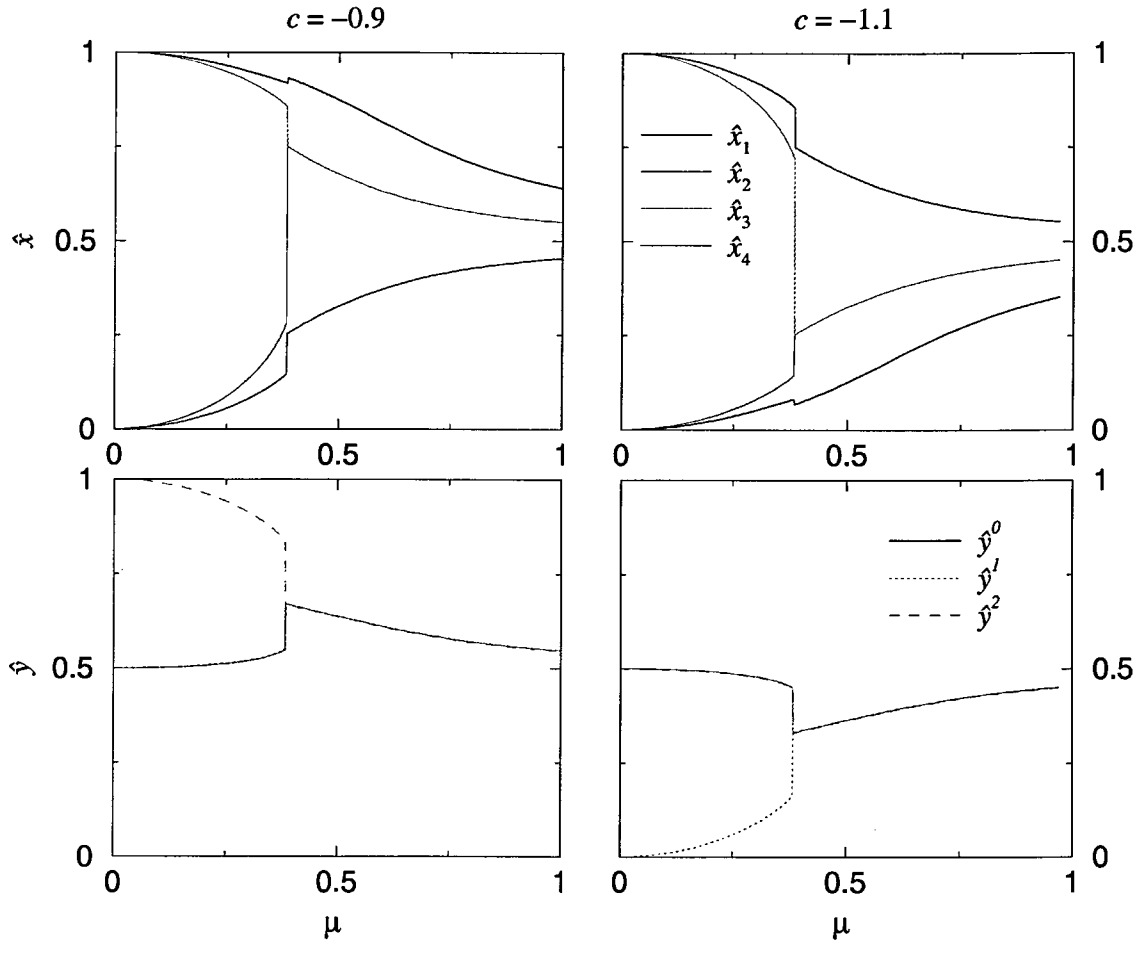

Figure 6.20: Ancestral mean partial distances $\hat{x}_v$ (top) and specific distances $\hat{y}^q$ (bottom) depending on mutation rates. The quadratic Hopfield-type fitness (6.18) with negative epistasis for three patterns and different values of c has been used. Results correspond to patterns for infinite sequence length.

Hopfield fitness (6.17) and symmetric patterns that correspond to an infinite sequence length, $X_v = 1/4$ for all v, which looks identical to the case of two patterns, cf. figure 6.16. The solutions for the different x_v all coincide. For small mutation rates $\mu < 1/2$, there are again two degenerate solutions for each x_v, which can be combined in multiple ways for the different x_v. At $\mu = 1/2$, there is a second order phase transition, which is a fitness and degradation threshold. The plot of the specific distances $\hat{y}^q$ (figure 6.19, right) reveals that for small mutation rates, the population is centred around one pattern, and random with respect to the others. In the case shown here, this one pattern is $\boldsymbol{\xi}^0$, as for all x_v, the lower of the two possible solutions has been chosen (cf. figure 6.19, left).

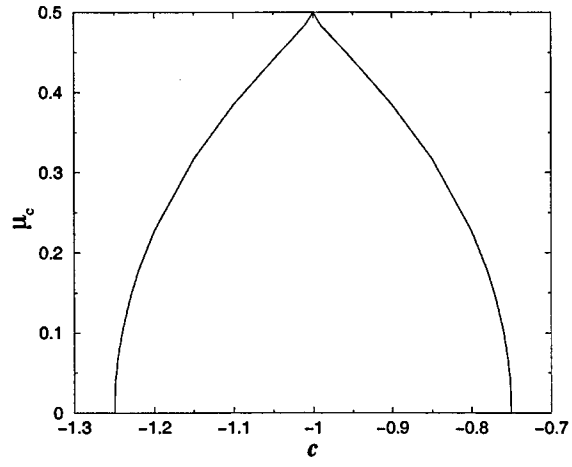

Figure 6.21: The critical mutation rate μ_c depending on the value of c. The quadratic Hopfield-type fitness (6.18) with negative epistasis for three patterns has been used. Results correspond to patterns for infinite sequence length.

Deviations from the original Hopfield fitness ($c \neq -1$). Figure 6.20 shows the ancestral mean partial distances $\hat{x}_v$ and specific distances $\hat{y}^q$ for the quadratic Hopfield-type fitness (6.18) with negative epistasis for different values of c. As c deviates from -1, the solutions for the x_v do not coincide, and the phase transition becomes a first order fitness threshold, at which all four partial distances $\hat{x}_v$ jump, but it is no more a degradation threshold. So contrary to the case of two patterns, where $\hat{x}_2$ is independent of c and the error threshold in $\hat{x}_1$ is smoothed out by c deviating from -1, here the threshold concerns all partial distances $\hat{x}_v$ and is sharpened to first order by $c \neq -1$.

For $c \neq -1$, the degeneracy between the patterns and their complements is lifted, and thus for mutation rates below the threshold, there are only three different solutions, correlated with the patterns for $c < -1$, and with their complements for $c > -1$. For clarity, only one of the solutions is shown in figure 6.20.

Furthermore, the critical mutation rate decreases with increasing $|c+1|$. The dependence of the critical mutation rate μ_c on the value of c is shown in figure 6.21. At $c = -5/4$ and $c = -3/4$, the critical mutation rate is $\mu_c = 0$, and for values of $c \notin [-5/4, -3/4]$, there are no error thresholds. This is an interesting result, as for all previously investigated Hopfield-type fitness functions (which to my knowledge are limited to the original Hopfield fitness and a Hopfield-type truncation selection), the existence of error thresholds

has been reported.

6.2.3.3 The asymmetric case simulating a finite sequence length

Figure 6.22 shows some typical cases of the ancestral mean partial and specific distances, $\hat{x}_v$ and $\hat{y}^q$, for the original Hopfield fitness (6.17) with three patterns, which are chosen randomly, with the X_v varying such as to simulate different sequence lengths. The six cases of patterns shown here are typical examples for the sequence lengths considered. In the case of long sequences ($N = 1000$), the deviations of the patterns from the infinite sequence limit $X_1 = X_2 = X_3 = X_4$ are small, and grow with decreasing sequence length. This "disorder" that is introduced into the system has the same effect as an asymmetric choice of patterns in the case of two patterns, such that the single critical mutation rate in the case of infinite sequence length is split up into two critical mutation rates, at each of which two of the x_v show threshold behaviour. For short sequence length ($N = 100$), it can be seen that particularly at the smaller of the critical mutation rates, the threshold is smoothed out.

In figure 6.23 the ancestral mean partial and specific distances, $\hat{x}_v$ and $\hat{y}^q$, corresponding to the same patterns as in figure 6.22 are shown for the quadratic Hopfield-type fitness (6.18) with negative epistasis and $c = -1.1$. For long sequence length, these look very similar to the results for infinite sequence length (cf. figure 6.20), showing clearly the single first order phase transition. For shorter sequence lengths, they become more and more smoothed out, such that at $N = 300$, roughly only every other pattern that was simulated shows an error threshold, whereas for $N = 100$, in the vast majority of cases, there is no threshold. Note that this effect is present even though the finite sequence length was only simulated by by choosing the patterns accordingly, so it is a feature of the model with asymmetric patterns.

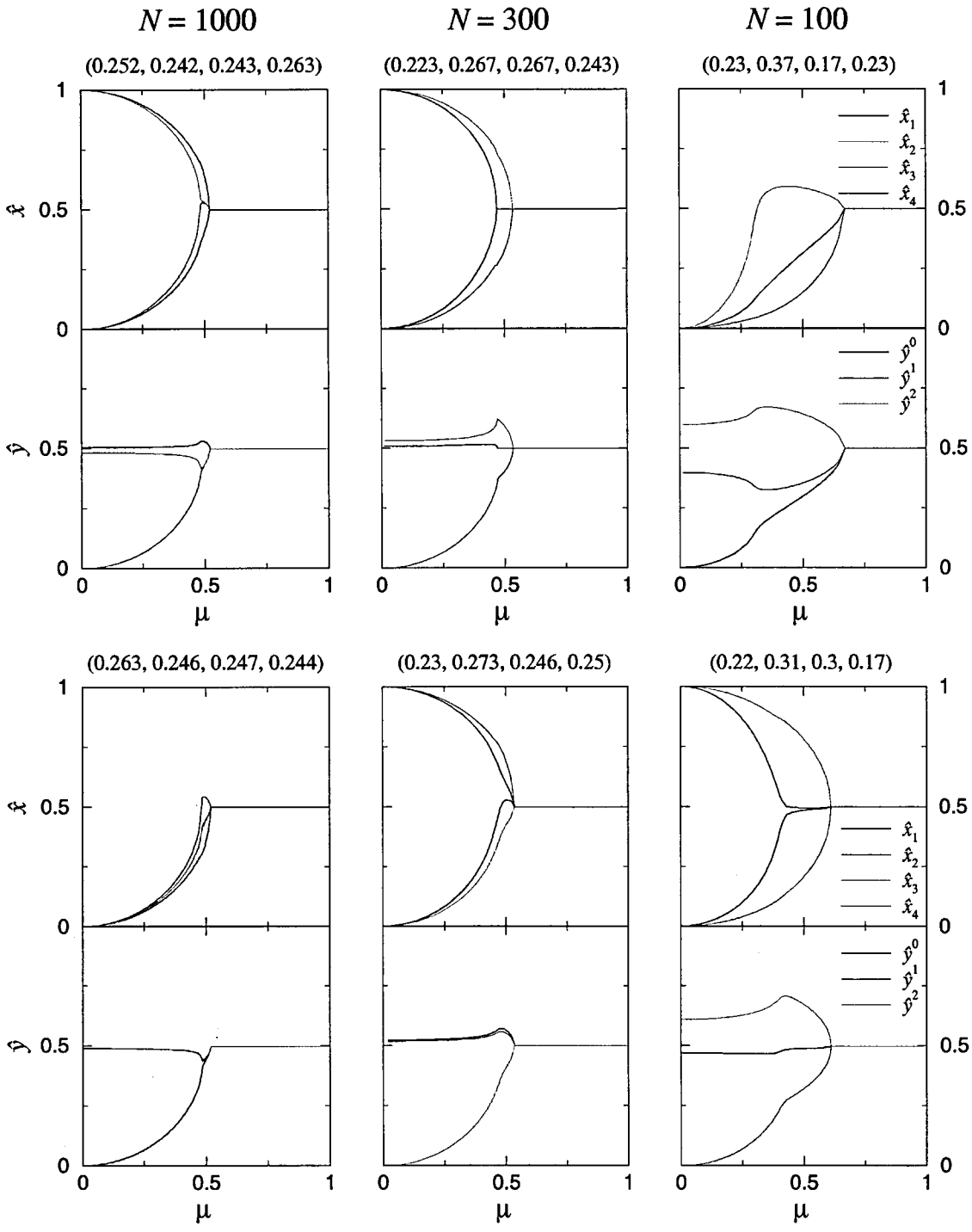

Figure 6.22: Ancestral mean partial and specific distances, $\hat{x}_v$ and $\hat{y}^q$, depending on mutation rates. The original Hopfield fitness (6.17) for three patterns has been used. Results correspond to two typical examples of random patterns chosen for sequences of lengths $N = 1000$ (left), $N = 300$ (middle) and $N = 100$ (right), specified at the top of each graph as (X_1, X_2, X_3, X_4).

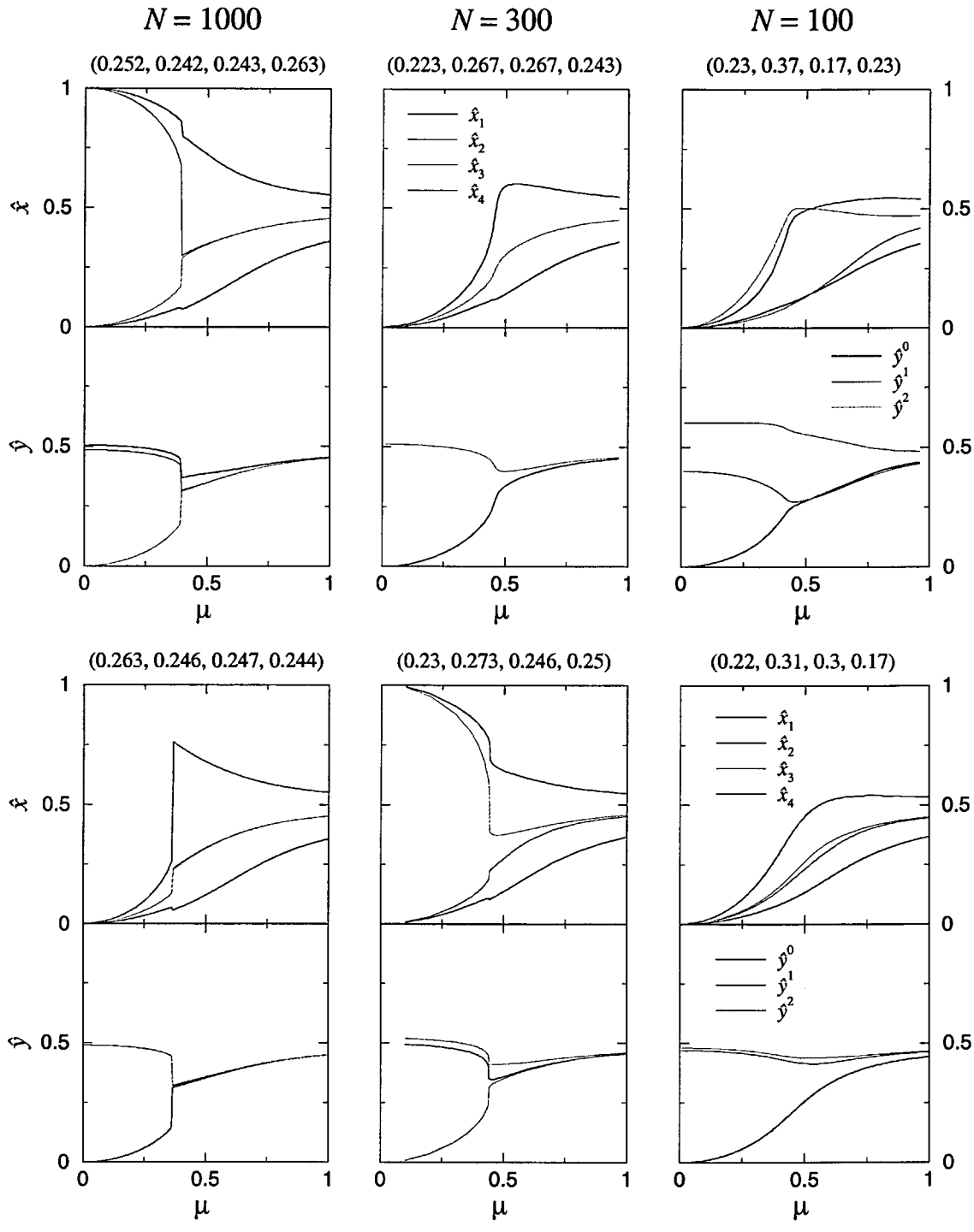

Figure 6.23: Ancestral mean partial and specific distances, $\hat{x}_v$ and $\hat{y}^q$, depending on mutation rates. The Hopfield-type fitness (6.18) with negative epistasis for three patterns and $c = -1.1$ has been used. Results correspond to the patterns used in figure 6.22.

$p=1$	symmetric patterns, $X_v = 1/2$	asymmetric patterns, $X_v \neq 1/2$
$c=-1$	one second order threshold for both x_v.	two second order thresholds, each for one of the x_v.
$c \neq -1$	one second order threshold in x_2, the x_1 threshold is smoothed out.	
$p=2$	symmetric patterns, $X_v = 1/4$	asymmetric patterns $X_v \neq 1/4$
$c=-1$	one second order threshold for all x_v.	two second order thresholds, each for two of the x_v.
$c \neq -1$	one first order threshold for all x_v if $c \in [-5/4, -3/4]$, no threshold otherwise.	up to one first order threshold for all x_v, but smoothed for shorter sequence lengths (stronger asymmetry).

Table 6.1: Summary of the results for Hopfield-type fitness (6.18) with two ($p = 1$) and three ($p = 2$) patterns.

6.2.4 Summary of the results for Hopfield-type fitness

In this section, the quadratic symmetric Hopfield-type fitness given in equation (6.18) with negative epistasis ($d = 1$) was investigated for the cases of two and three patterns. The results for the original Hopfield fitness ($c = -1$) have been compared with those for the generalised Hopfield-type fitness ($c \neq 1$). Furthermore, both symmetric patterns ($X_v = 2^{-(p+1)}$ for all v), corresponding to an infinite sequence length, and asymmetric patterns ($X_v \neq 2^{-(p+1)}$), simulating a finite sequence length, were considered. For two patterns, an analytical treatment was possible, making all values of the X_v accessible, whereas the case of three patterns was treated numerically due to the larger number of variables. This means that apart from the symmetric choice of pattern ($X_v = 1/4$), only some asymmetric combinations with $X_v \neq 1/4$ were investigated, some typical examples of which are shown in section 6.2.3.

The results obtained for Hopfield-type fitness with two and three patterns are summarised in table 6.1. It shows that for the original Hopfield fitness ($c = -1$), the systems with two and three patterns behave very similar, showing two thresholds for asymmetric patterns, which coincide in the case of symmetric patterns. However, with a Hopfield-

type fitness with $c \neq -1$, the results for two and three patterns are quite different. For two patterns with $c \neq -1$, irrespective of whether symmetric or asymmetric patterns are chosen, the threshold in x_2 is unaffected, because it cancels out in the linear part of the fitness, which is tuned by the parameter c; and therefore it stays a second order threshold, but the threshold in x_1 is smoothed out when c deviates from -1.

For three patterns that are symmetric, the single second order threshold in all variables that exists in the original Hopfield fitness is sharpened to first order when $c \neq -1$, going in line with a decrease in the critical mutation rate μ_c. However, if c deviates too much, i.e., $c \notin [-5/4, -3/4]$, there is no threshold as the critical mutation rate crosses $\mu_c = 0$ and becomes negative. The effect of finite sequence length, obtained for asymmetric patterns, in this case is only to smooth out the single first order threshold, if the asymmetry (or disorder) becomes too strong, which happens for shorter sequence lengths.

The evaluation the Hopfield-type fitness was limited to the cases of two and three patterns, simply because an increase in the number of patterns makes the evaluation more complex. However, the Hopfield-type fitness was chosen as a potentially realistic fitness because of its ruggedness that can be tuned by the number of patterns chosen. The simple cases considered here probably do not show as high a degree of ruggedness as one would expect for realistic fitness functions. However, some of the features observed here might also occur for a higher number of patterns. It would thus be very interesting to investigate the system further with respect to a larger number of patterns, and to establish which of the features described for the two and three pattern case generalise to any number of patterns, and which may depend on whether the number of patterns is odd or even.

Furthermore, the concept of partitioning the set of sites into subsets, which was introduced to analyse the Hopfield-type fitness, is very interesting. One could imagine a different interpretation for this by classifying sites according to the selection strength they evolve under. Some of the behaviour identified for the Hopfield system could also occur in such a setting: At intermediate mutation rates there partially ordered phases could exist, such that sites that evolve under weak selection have passed their error threshold and the

population is in a phase that is disordered with respect to these sites, whereas at sites that are subject to strong selection the order is still maintained.

CHAPTER 7

Conclusion

This thesis has been concerned with the investigation of a deterministic mutation-selection model in the sequence space approach. The particular model chosen in this work is the para-muse model, where time is treated as a continuous quantity, modelling generations as overlapping, in contrast to some other works drawing on the connection to statistical mechanics [Leu86, Tar92, Gal97]. Furthermore, in the para-muse model, mutation and selection are treated as two decoupled processes, going on in parallel. This differs from the original quasispecies model [Eig71], one of the first to use the sequence space approach in population genetics. However, it has been shown that these differences in models have relatively little impact on the outcomes; generally, the results obtained in any mutation–selection model in the sequence space approach are similar [WBS95, Bür98].

Important observables in mutation–selection models are means of quantities such as fitness and genotype with respect to the population and ancestral distributions. The population distribution $\boldsymbol{p}(t)$ gives the fraction of individuals of each type; in equilibrium it is given by the right PF-eigenvector $\boldsymbol{p}$ of the time-evolution operator $\boldsymbol{H}$. As $\boldsymbol{p}$ is normalised such that $\sum_i p_i(t) = 1$, the population mean of a quantity o, given by $\bar{o}(t) = \sum_i o_i p_i(t)$ is normalised with respect to an L_1-norm.

The ancestral distribution $\boldsymbol{a}(\tau, t)$, a concept introduced only recently [HRWB02], is the population distribution at time t, but weighted with the number of offspring stemming from each individual at a later time $t + \tau$. In equilibrium, the ancestral distribution is given by both the left and right PF-eigenvectors $\boldsymbol{z}$ and $\boldsymbol{p}$ of $\boldsymbol{H}$, $a_i = z_i p_i$, such that the ancestral mean of a quantity o, $\hat{o}(\tau, t) = \sum_i o_i a_i(\tau, t)$ is normalised with respect to an L_2-norm.

In the sequence space approach, genotypes are identified with sequences written in two- or four-letter alphabets to mimic the genetic code. The mutation process is modelled as a Markov process, with one mutation rate in the two-state case and three mutation rates according to the Kimura 3ST mutation scheme in the four-state case.

Modelling selection appropriately is however far more difficult. One very common approach is to consider permutation-invariant fitness functions, which only depend on the number of mutations a sequence carries compared to the wild-type, i.e., the Hamming distance d_H to the wild-type, not on their location within the sequence. Although this is clearly an unrealistic assumption, it is a good starting point. Approaches using non-permutation-invariant fitness functions often draw on Hamiltonians of spin glass models in physics. In this thesis, fitness functions of the Hopfield type as an example of a spin glass fitness have been investigated alongside permutation-invariant fitness functions.

One problem in the sequence space approach is the large number of types, as there are 2^N or 4^N different sequences of length N in the two- or four-state model, respectively. This problem was addressed in chapter 4 by lumping together sequences with the same Hamming distance into classes, creating a new coarser process acting on the classes rather than on individual types. To this end, the concept of the scalar Hamming distance, which is relevant in the two-state case with permutation-invariant fitness, was generalised to a vector quantity, used in the four-state case with permutation-invariant fitness and for models with Hopfield-type fitness.

A major advance for determining the population mean fitness in equilibrium was made in [HRWB02], where a simple scalar maximum principle was derived, applicable to a two-state model with permutation-invariant fitness. In chapter 5, this scalar maximum principle was generalised to a vectorial one for the four-state and Hopfield-type cases. The maximiser can be interpreted as the ancestral mean genotype $\hat{\boldsymbol{x}}$, lending importance to the ancestral mean.

In chapter 6, the maximum principles derived in chapter 5 were used to investigate the phenomenon of the error threshold. First discovered in [Eig71], the error threshold

describes the transition of a population with a distribution that is well localised in sequence space to a less localised distribution or ultimately the equidistribution in sequence space that occurs at a critical mutation rate. These error thresholds can be detected with the maximum principle, because the delocalisation of the population distribution manifests itself as a jump (or at least an infinite derivative with respect to the mutation rate) of the ancestral mean genotype $\hat{x}$, the maximiser. Not all fitness functions give rise to error thresholds, and as the error thresholds were first described for a model with highly unrealistic fitness function, it has been argued that they might be an artefact of this rather than a biologically relevant phenomenon. It is therefore clearly necessary to investigate other fitness functions with respect to this phenomenon.

For permutation-invariant fitness functions, the four-state model was investigated. Mutation was modelled using the Kimura 2 parameter model, a simplification of the full Kimura 3ST mutation scheme, using only two different mutation rates μ and μ_2. There are no error thresholds for linear fitness functions, and therefore the focus was on slightly more complex cases such as quadratic fitness functions and truncation selection. As far as quadratic fitness functions are concerned, it transpires that only those with negative epistasis, i.e., positive curvature, give rise to error thresholds.

Apart from the ordered and disordered phases, which also occur in two-state models, a partially ordered phase was observed [HWB01], which differentiates only between purines and pyrimidines. This phase directly hinges on the multidimensional mutational distance that is employed in the four-state model. However, as this analysis has shown, it is also dependent on a certain symmetry of the fitness function which implies that transition-type mutations have no effect on the fitness or, in other words, that the fitness function is symmetric with respect to the fraction of wild-type sites x_0 and the fraction of transition-type sites x_2. Whereas this symmetry comes naturally in the equivalent physical systems as 'spin-flip symmetry', it seems rather artificial in the biological setting. If one considers only the third-codon positions in coding regions of the DNA though, the vast majority of the transition-type mutations is silent, i.e., they do not change the amino acid that

is incorporated in the protein in question, with the result that these mutations influence the fitness only minimally. Consequently, one would not expect to find a global partially ordered phase in any organism, but among the sites at third-codon positions, partially ordered phases might occur.

Quadratic Hopfield-type fitness functions have been investigated for the two-state model with two and three patterns. The results for both differ if the fitness deviates from the original Hopfield fitness. This raises the question which of the differences observed can be attributed to whether the number of patterns is even or odd and therefore can be generalised to a higher number of patterns, and which may be particular to the two- or three-pattern case, respectively. Furthermore, in the original Hopfield fitness, error thresholds were observed for all choices of patterns. This is not true for generalised Hopfield-type fitness. For instance, for a Hopfield-type fitness with positive epistasis no thresholds were observed, going in line with the results for permutation-invariant fitness. But also for Hopfield-type fitness functions with negative epistasis, there are not necessarily any thresholds, if the fitness deviates too much from the original Hopfield fitness, challenging the commonly held notion that more complex fitness functions all tend to display error threshold behaviour. The complexity and ruggedness of the original Hopfield fitness have been investigated [AGS85a, AGS85b] and found to be good candidates for realistic fitness functions [Leu87, Tar92]. However, these results do not necessarily transfer to the generalised Hopfield-type fitness functions, and therefore it would be very useful to study these properties of generalised Hopfield-type fitness functions to analyse which of these factors are responsible for generating the thresholds.

The main results of this thesis are the derivation of the maximum principles obtained in chapter 5 and their application to several examples pursued in chapter 6. The maximum principle for a four-state model with permutation-invariant fitness has already been published in [GG04b], its application to quadratic fitness and truncation selection is published in [GG04a]. The results concerning the Hopfield-type fitness have not been published yet.

The mutation–selection models considered here describe biological evolution clearly

in a rather simplistic way, neglecting a number of biologically relevant factors such as, for instance, recombination or migration, and possibly most importantly, genetic drift. Furthermore, only the equilibrium properties were considered. However, a solid understanding of the mutation-selection balance is a prerequisite to investigate the approach to equilibrium, and a good starting point for incorporation of other evolutionary factors.

7.1 Future work

Using the deterministic model as set out in this thesis, one could try to generalise the criteria for fitness functions to give rise to error thresholds as obtained in [HRWB02] for a two-state model with permutation-invariant fitness to models with multidimensional mutational distances such as the four-state model or a model with Hopfield-type fitness. Furthermore, it would be interesting to investigate the Hopfield-type fitness with respect to a larger number of patterns.

In order to make the model biologically more realistic, it would be very interesting include genetic drift and consider a stochastic model such as a Moran model [Mor58, Ewe04] with selection.

References

[AGS85a] D. J. Amit, H. Gutfreund, and H. Sompolinsky. Spin-glass models of neural networks. *Physical Review A*, 32(2):1007–1018, 1985.

[AGS85b] D. J. Amit, H. Gutfreund, and H. Sompolinsky. Storing infinite numbers of patterns in a spin-glass model of neural networks. *Physical Review Letters*, 55(14):1530–1533, 1985.

[Aki79] E. Akin. *The Geometry of Population Genetics*, volume 31 of *Lecture Notes in Biomathematics*. Springer, New York, 1979.

[And83] P. W. Anderson. Suggested model for prebiotic evolution: The use of chaos. *Proceedings of the National Academy of Sciences of the USA*, 80(11):3386–3390, 1983.

[Baa95] E. Baake. Diploid models on sequence space. *Journal of Biological Systems*, 3(2):343–349, 1995.

[BABG03] G. Ben-Arous, A. Bovier, and V. Gayrard. Glauber dynamics of the random energy model. 1. metastable motion on the extreme states. *Communications in Mathematical Physics*, 235(3):379–425, 2003.

[BB00] F. Bagnoli and M. Bezzi. An evolutionary model for simple ecosystems. In D. Stauffer, editor, *Annual Reviews of Computational Physics VII*, page 265. World Scientific, Singapore, 2000.

[BBBK05] E. Baake, M. Baake, A. Bovier, and M. Klein. An asymptotic maximum principle for essentially linear evolution models. *Journal of Mathematical Biology*, 50(1):83–114, 2005.

[BBN96] M. C. Boerlijst, S. Bonhoeffer, and M. A. Nowak. Viral quasi-species and recombination. *Proceedings of the Royal Society of London, Series B*, 263(1376):1577–1584, 1996.

[BBW97] E. Baake, M. Baake, and H. Wagner. Ising quantum chain is equivalent to a model of biological evolution. *Physical Review Letters*, 78(3):559–562, 1997. Erratum, Physical Review Letters 79 (1997), 1782.

[BBW98] E. Baake, M. Baake, and H. Wagner. Quantum mechanics versus classical probability in biological evolution. *Physical Review E*, 57(1):1191–1192, January 1998.

[BG00] E. Baake and W. Gabriel. Biological evolution through mutation, selection, and drift: An introductory review. In D. Stauffer, editor, *Annual Reviews of Computational Physics VII*, pages 203–264. World Scientific, Singapore, 2000.

[BS93] S. Bonhoeffer and P. F. Stadler. Error thresholds on correlated fitness landscapes. *Journal of Theoretical Biology*, 164(3):359–372, 1993.

[Bür98] R. Bürger. Mathematical properties of mutation-selection models. *Genetica*, 103:279–298, 1998.

[Bür00] R. Bürger. *The Mathematical Theory of Selection, Recombination, and Mutation.* Wiley, Chichester, 2000.

[BW01] E. Baake and H. Wagner. Mutation-selection models solved exactly with methods of statistical mechanics. *Genetical Research*, 78:93–117, 2001.

[Cas01] H. Caswell. *Matrix Population Models.* Sinauer, Sunderland, 2nd edition, 2001.

[CAW02] P. R. A. Campos, C. Adami, and C. O. Wilke. Optimal adaptive performance and delocalization in NK fitness landscapes. *Physica A*, 304(3–4):495–506, 2002.

[CCA01] S. Crotty, C. E. Cameron, and R. Andino. RNA virus error catastrophe: Direct molecular test by using ribavirin. *Proceedings of the National Academy of Sciences of the USA*, 98(12):6895–6900, 2001.

[CE92] M. E. Cochran and S. Ellner. Simple methods for calculating age-based life history parameters for stage-structured populations. *Ecological Monographs*, 62(3):345–364, 1992.

[Cha88] L. Chao. Evolution of sex in RNA viruses. *Journal of Theoretical Biology*, 133(1):99–112, 1988.

[CK70] J. F. Crow and M. Kimura. *An Introduction to Population Genetics Theory.* Harper & Row, New York, 1970.

[CL04] J. H. Chuang and H. Li. Functional bias and spatial organization of genes in mutational hot and cold regions in the human genome. *PLoS Biology*, 2(2):0253–0263, 2004. (Preprint q-bio.GN/0402006).

[DES+96] E. Domingo, C. Escarmis, N. Sevilla, A. Moya, S. F. Elena, J. Quer, I. S. Novella, and J. J. Holland. Basic concepts in RNA virus evolution. *FASEB Journal*, 10(8):859–864, 1996.

[DH88] E. Domingo and J. J. Holland. High error rates, population equilibrium, and evolution of RNA replication systems. In E. Domingo, editor, *RNA Genetics*, volume 3, page 3. CRC Press, Boca Raton, 1988.

[DH97] E. Domingo and J. J. Holland. RNA virus mutations and fitness for survival. *Annual Review of Microbiology*, 51:151–178, 1997.

[EB88] M. Eigen and C. K. Biebricher. Sequence space and quasispecies evolution. In E. Domingo, editor, *RNA Genetics*, volume 3, pages 211–245. CRC Press, Boca Raton, 1988.

[Eig71] M. Eigen. Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften*, 58(10):465–523, 1971.

[Eig93] M. Eigen. Viral quasispecies. *Scientific American*, 269(1):42–49, 1993.

[EK86] S. N. Ethier and T. G. Kurtz. *Markov Processes - Characterization and Convergence.* Wiley, New York, 1986.

[EMS89] M. Eigen, J. McCaskill, and P. Schuster. The molecular quasi-species. *Advances in Chemical Physics*, 75:149–263, 1989.

[Ewe79] W. J. Ewens. *Mathematical Population Genetics.* Springer, Berlin, 1979.

[Ewe04] W. J. Ewens. *Mathematical Population Genetics.* Springer, New York, 2nd edition, 2004.

[FFHS00] C. Flamm, W. Fontana, I. L. Hofacker, and P. Schuster. RNA folding at elementary step resolution. *RNA*, 6(3):325–338, 2000.

[Fis30] R. A. Fisher. *The genetical theory of natural selection.* Oxford University Press, 1930.

[FP97] S. Franz and L. Peliti. Error threshold in simple landscapes. *Journal of Physics A*, 30(13):4481–4487, 1997.

[FPS93] S. Franz, L. Peliti, and M. Sellitto. An evolutionary version of the random energy model. *Journal of Physics A*, 26(23):L1195–L1199, 1993.

[FS98a] W. Fontana and P. Schuster. Continuity in evolution: On the nature of transitions. *Science*, 280:1451–1455, 1998.

[FS98b] W. Fontana and P. Schuster. Shaping space: the possible and the attainable in RNA genotype–phenotype mapping. *Journal of Theoretical Biology*, 194(4):491–515, 1998.

[Gal97] S. Galluccio. Exact solution of the quasispecies model in a sharply peaked fitness landscape. *Physical Review E*, 56(4):4526–4539, 1997.

[Gay92] V. Gayrard. Thermodynamic limit of the q-state Potts–Hopfield model with infinitely many patterns. *Journal of Statistical Physics*, 68(5/6):977–1011, 1992.

[GdJ93] S. Gavrilets and G. de Jong. Pleiotropic models of polygenic variation, stabilizing selection, and epistasis. *Genetics*, 134(2):609–625, 1993.

[GG04a] T. Garske and U. Grimm. Maximum principle and mutation thresholds for four-letter sequence evolution. *Journal of Statistical Mechanics: Theory and Experiment*, P07007, 2004. (Preprint q-bio.PE/0406041).

[GG04b] T. Garske and U. Grimm. A maximum principle for the mutation–selection equilibrium of nucleotide sequences. *Bulletin of Mathematical Biology*, 66(3):397–421, 2004. (Preprint physics/0303053).

[GGZ96] S. Galluccio, R. Graber, and Y.-C. Zhang. Diffusion on a hypercubic lattice with pinning potential: exact results for the error-catastrophe problem in biological evolution. *Journal of Physics A*, 29(10):L249–L255, 1996.

[GH02] U. Gerland and T. Hwa. On the selection and evolution of regulatory DNA motifs. *Journal of Molecular Evolution*, 55:386–400, 2002.

[Gil91] J. H. Gillespie. *The Causes of Molecular Evolution.* Oxford University Press, New York, 1991.

[Hal27] J. B. S. Haldane. A mathematical theory of natural and artificial selection. part v: Selection and mutation. *Proceedings of the Cambridge Philosophical Society*, 23:838–844, 1927.

[Ham50] R. W. Hamming. Error detecting and error correcting codes. *The Bell System Technical Journal*, 26(2):147–160, 1950.

[HDdlTS90] J. J. Holland, E. Domingo, J. C. de la Torre, and D. A. Steinhauer. Mutation frequencies at defined single codon sites in vesicular stromatitis-virus and poliovirus can be increased only slightly by chemical mutagenesis. *Journal of Virology*, 64(8):3960–3962, 1990.

[HFS+94] I. L. Hofacker, W. Fontana, P. F. Stadler, L. S. Bonhoeffer, M. Tacker, and P. Schuster. Fast folding and comparison of RNA secondary structures. *Monatshefte für Chemie*, 125:167–188, 1994. The Vienna RNA software package is available at `http://www.tbi.univie.ac.at/~ivo/RNA`.

[Hig94] P. Higgs. Error thresholds and stationary mutant distributions in multilocus diploid genetics models. *Genetical Research Cambridge*, 63(1):63–78, 1994.

[Hig00] P. G. Higgs. RNA secondary structure: Physical and computational aspects. *Quarterly Reviews of Biophysics*, 33(3):199–253, 2000.

[Hop82] J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the USA*, 79(8):2554–2558, 1982.

[HRWB02] J. Hermisson, O. Redner, H. Wagner, and E. Baake. Mutation selection balance: Ancestry, load, and maximum principle. *Theoretical Population Biology*, 62:9–46, 2002.

[HSF96] M. A. Huynen, P. F. Stadler, and W. Fontana. Smoothness within ruggedness: the role of neutrality in adaptation. *Proceedings of the National Academy of Sciences of the USA*, 93(1):397–401, 1996.

[HW01] T. F. Hansen and G. P. Wagner. Epistasis and the mutation load: A measurement-theoretical approach. *Genetics*, 158(1):477–485, 2001.

[HWB01] J. Hermisson, H. Wagner, and M. Baake. Four-state quantum chain as a model of sequence evolution. *Journal of Statistical Physics*, 102(1/2):315–343, 2001.

[JC69] T. H. Jukes and C. R. Cantor. Evolution of protein molecules. In H. N. Munro, editor, *Mammalian Protein Metabolism*, pages 21–132. Academic Press, New York, 1969.

[Kar66] S. Karlin. *A First Course in Stochastic Processes.* Academic Press, New York, 1966.

[KB03] L. Katz and C. B. Burge. Widespread selection for local RNA secondary structure in coding regions of bacterial genes. *Genome Research*, 13(9):2042–2051, 2003.

[Kim69] M. Kimura. The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. *Genetics*, 61:893–903, 1969.

[Kim80] M. Kimura. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, 16:111–120, 1980.

[Kim81] M. Kimura. Estimation of evolutionary distances between homologous nucleotide sequences. *Proceedings of the National Academy of Sciences of the USA*, 78(1):454–458, 1981.

[Kim83] M. Kimura. *The neutral theory of molecular evolution.* Cambridge University Press, Cambridge, 1983.

[Kim87] M. Kimura. Molecular evolutionary clock and the neutral theory. *Journal of Molecular Evolution*, 26(1–2):24–33, 1987.

[KL87] S. Kauffman and S. Levin. Towards a general theory of adaptive walks on rugged landscapes. *Journal of Theoretical Biology*, 128:11–45, 1987.

[Kon88] A. S. Kondrashov. Deleterious mutations and the evolution of sexual reproduction. *Nature*, 336:435–440, 1988.

[Kon98] A. S. Kondrashov. Measuring spontaneous deleterious mutation process. *Genetica*, 102/103:183–197, 1998.

[KS60] J. G. Kemeny and J. L. Snell. *Finite Markov Chains.* Van Nostrand Reinhold Company, New York, 1960.

[LEK+99] L. A. Loeb, J. M. Essigmann, F. Kazazi, J. Zhang, K. D. Rose, and J. I. Mullins. Lethal mutagenesis of HIV with mutagenic nucleoside analogs. *Proceedings of the National Academy of Sciences of the USA*, 96:1492–1497, 1999.

[Leu86] I. Leuthäusser. An exact correspondence between Eigen's evolution model and a two-dimensional Ising system. *Journal of Chemical Physics*, 84(3):1884–1885, 1986.

[Leu87] I. Leuthäusser. Statistical mechanics of Eigen's evolution model. *Journal of Statistical Physics*, 48(1/2):343–360, 1987.

[LTP03] M. Lässig, F. Tria, and L. Peliti. Evolutionary games and quasispecies. *Europhysics Letters*, 62(3):446–451, 2003.

[May83] J. Maynard Smith. Models of evolution. *Proceedings of the Royal Society of London, Series B*, 219(1216):315–325, 1983.

[Mor58] P. A. P. Moran. Random processes in genetics. *Proceedings of the Cambridge Philosophical Society*, 54:60, 1958.

[MPV87] M. Mézard, G. Parisi, and M. A. Virasoro. *Spin Glass Theory and Beyond.* World Scientific, Singapore, 1987.

[MS95] J. Maynard Smith and E. Szathmáry. *The Major Transitions in Evolution.* Freeman, Oxford, 1995.

[Muk64] T. Mukai. The genetic structure of natural populations of *drosophila melanogaster*. I. Spontaneous mutation rate of polygenes controlling viability. *Genetics*, 50:1–19, 1964.

[Muk69] T. Mukai. The genetic structure of natural populations of *drosophila melanogaster*. VII. Synergistic interaction of spontaneous mutant polygenes controlling viability. *Genetics*, 61:749–761, 1969.

[NS89] M. Nowak and P. Schuster. Error thresholds of replication in finite populations. Mutation frequencies and the onset of Muller's ratchet. *Journal of Theoretical Biology*, 137(4):375–395, 1989.

[OK73] T. Ohta and M. Kimura. A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. *Genetical Research*, 22:201–204, 1973.

[Oth93] T. Otha. An examination of the generation-time effect on molecular evolution. *Proceedings of the National Academy of Sciences of the USA*, 90(22):10676–10680, 1993.

[Pel02] L. Peliti. Quasispecies evolution in general mean-field landscapes. *Europhysics Letters*, 57(5):745–751, 2002.

[Per07] O. Perron. Zur Theorie der Matrices. *Mathematische Annalen*, 64:248–263, 1907.

[Red03] O. Redner. Discretization of continuum-of-alleles models. *Preprint*, 2003. (Preprint math.SP/0301024).

[RFS01] C. Reidys, C. V. Forst, and P. Schuster. Replication and mutation on neutral networks. *Bulletin of Mathematical Biology*, 63(1):57–94, 2001.

[RS02] C. M. Reidys and P. F. Stadler. Combinatorial landscapes. *SIAM Review*, 44(1):3–54, 2002.

[Rum87] D. S. Rumschitzky. Spectral properties of Eigen's evolution matrices. *Journal of Mathematical Biology*, 24:667–680, 1987.

[SDLD00] S. Sierra, M. Dávila, P. R. Lowenstein, and E. Domingo. Response of foot-and-mouth disease virus to increased mutagenesis: Influence of viral load and fitness in loss of infectivity. *Journal of Virology*, 74(18):8316–8323, 2000.

[SK75] D. Sherrington and S. Kirkpatrick. Solvable model of a spin glass. *Physical Review Letters*, 35(26):1792–1796, 1975.

[SOWH96] D. L. Swofford, G. J. Olsen, P. J. Waddell, and D. M. Hillis. Phylogenetic inference. In David M. Hillis, Craig Moritz, and Barbara K. Mable, editors, *Molecular Systematics*, pages 407–514. Sinauer, Sunderland, 1996.

[Tal03] M. Talagrand. *Spin Glasses: A Challenge for Mathematicians.* Springer, Berlin, 2003.

[Tar92] P. Tarazona. Error thresholds for molecular quasispecies as phase transitions: From simple landscapes to spin-glass models. *Physical Review A*, 45(8):6038–6050, 1992.

[TLK96] L. S. Tsimring, H. Levine, and D. A. Kessler. RNA virus evolution via a fitness-space model. *Physical Review Letters*, 76(23):4440–4443, 1996.

[TM74] C. J. Thompson and J. L. McBride. On Eigen's theory of the self-organization of matter and the evolution of biological macromolecules. *Mathematical Biosciences*, 21(1–2):127–142, 1974.

[vHB86] P. H. von Hippel and O. G. Berg. On the specificity of DNA–protein interactions. *Proceedings of the National Academy of Sciences of the USA*, 83:1608–1612, 1986.

[vL82] J. H. van Lint. *Introduction to Coding Theory.* Springer, Berlin, 1982.

[WA01] C. O. Wilke and C. Adami. Interaction between directional epistasis and average mutational effects. *Proceedings of the Royal Society London, Series B*, 268(1475):1469–1474, 2001.

[Wag98] H. Wagner. *Biologische Sequenzraummodelle und Statistische Mechanik.* PhD thesis, Universität Tübingen, 1998.

[WBS95] T. Wiehe, E. Baake, and P. Schuster. Error propagation in reproduction of diploid organisms. A case study on single peaked landscapes. *Journal of Theoretical Biology*, 177(1):1–15, 1995.

[WC53] J. D. Watson and F. H. C. Crick. Molecular structure of nucleic acids. A structure for desoxyribose nucleic acid. *Nature*, 171:737–738, 1953.

[WG01] S. Whelan and N. Goldman. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Molecular Biology and Evolution*, 18(5):691–699, 2001.

[Whi76] P. Whittle. *Probability.* Wiley, London, 1976.

[Wie97] T. Wiehe. Model dependency of error thresholds: The role of fitness functions and contrasts between finite and infinite sites models. *Genetical Research Cambridge*, 69:127–136, 1997.

[WLA03] C. O. Wilke, R. E. Lenski, and C. Adami. Compensatory mutations cause excess of antagonistic epistasis in RNA secondary structure folding. *BMC Evolutionary Biology*, 3(3), 2003.

[WLG01] S. Whelan, P. Liò, and N. Goldman. Molecular phylogenetics: state-of-the-art methods for looking into the past. *Trends in Genetics*, 17(5):262–272, 2001.

[Yan96] Z. Yang. Among-site rate variation and its impact on phylogenetic analyses. *Trends in Ecology and Evolution*, 11(9):367–372, 1996.